A Publication of the American Psychological Association

Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles

Communality and Stability of Meaning in Clinical Case Description 389
Erika Chance, Jack Arnold, and Sybil Tyrrell

Ego Involvement: A Critical Examination of Some Methodological Issues 407
Eva Dreikurs Ferguson

Parental Attitudes Associated with Social Deviance in Preadolescent Boys 418
C. L. Winder and Lucy Rau

Preference for Delayed Reward as a Function of Age, Intelligence, and Length of Delay Interval 425
Walter Mischel and Ralph Metzner

Anxiety and Avoidance Behavior in Schizophrenics in Response to Parental Figures 432
James C. Baxter and Joseph Becker

Response to Conflict as Determined by Prior Exposure to Conflict 438
Leonard Worell

Critique and Notes

The Effects of Three Types of Aggression on Vascular Processes 446
Jack E. Hokanson and Michael Burgess

The Effect of Affective Tone on the Verbal Behavior of Process and Reactive Schizophrenics 450
Alan S. DeWolfe

Number of Verbal Interrelationships as a Determinant of Group Size ... 456
George F. Castore

Relationship between Probability of Item Endorsement and Social Desirability Scale Value for High and Low Groups on Edwards' *SD* Scale 458
Allen L. Edwards and Jerald N. Walker

This is the last issue of Volume 64.
Title page and Contents for the Volume appear herein.

Volume 64 JUNE 1962 Number 6

Volume 64 June 1962 Number 6

Journal of Abnormal and Social Psychology

Editor: DANIEL KATZ, *University of Michigan*

Associate Editor: ROBERT B. ZAJONC, *University of Michigan*

CONSULTING EDITORS

JOSEPH B. ADELSON
University of Michigan

GORDON W. ALLPORT
Harvard University

NORMAN H. ANDERSON
University of California, Los Angeles

LEONARD BERKOWITZ
University of Wisconsin

JACK W. BREHM
Duke University

URIE BRONFENBRENNER
Cornell University

DONALD R. BROWN
Bryn Mawr College

DONALD T. CAMPBELL
Northwestern University

LAUNOR F. CARTER
System Development Corporation

DWIGHT W. CHAPMAN, JR.
Vassar College

MORTON DEUTSCH
Bell Telephone Laboratories

SEYMOUR FESHBACH
University of Pennsylvania

DONALD W. FISKE
University of Chicago

NATHANIEL L. GAGE
University of Illinois

NICHOLAS HOBBS
Peabody College for Teachers

MURRAY HORWITZ
New York University

IRVING L. JANIS
Yale University

HAROLD H. KELLEY
University of California, Los Angeles

HERBERT C. KELMAN
Harvard University

DANIEL J. LEVINSON
Harvard Medical School

DAVID C. MCCLELLAND
Harvard University

ROBERT B. MACLEOD
Cornell University

HAROLD L. RAUSH
National Institute of Mental Health

NEVITT SANFORD
Stanford University

IRVING SARNOFF
Western Reserve University

WILLIAM A. SCOTT
University of Colorado

JANET TAYLOR SPENCE
Veterans Administration Hospital, Iowa City

JOSEPH VEROFF
University of Michigan

JUSTIN L. WEISS
Massachusetts Mental Health Center

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued monthly, two volumes per year. The subscription price per volume is $10.00, or $20.00 annually (foreign, $20.50). Single issues are $2.00.

Original manuscripts should be sent to the Editor:

DANIEL KATZ
University of Michigan
Box 1248
Ann Arbor, Michigan

Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Acting Production Manager

JUDY SMITH
Editorial Assistant

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street, N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth Street N.W., Washington 6, D. C.

Second-class postage paid at Lancaster, Pa., and at additional mailing offices

Copyright © 1962 by the American Psychological Association, Inc.

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 389–406

COMMUNALITY AND STABILITY OF MEANING IN CLINICAL CASE DESCRIPTION[1]

ERIKA CHANCE, JACK ARNOLD, AND SYBIL TYRRELL

Department of Psychiatry, Mount Zion Hospital, San Francisco

One major difficulty in communication among clinicians is that much depends on the point of view of the observer and on his peculiar emphasis of selected aspects of behavior as crucial variables in case description. Most clinical investigations employ variables for describing behavior without questioning whether the variables mean the same thing when applied by different observers or when applied to different patients. In many such studies it is commonly assumed that each of the variables selected corresponds to a unitary concept that is used with more or less emphasis by various kinds of observers (observer bias) and is found to be more or less characteristic of various kinds of patients (differences among patients). These studies focus on differences in score level among observer or patient populations.

In our earlier work (Chance, 1959; Chance & Arnold, 1960) with a system of variables for classifying interpersonal experiences (see Figure 1), we assumed, first, that the 20 variables which compose the system had the same meanings for a variety of observers and that the meaning of each variable was stable regardless of the patient being described. Inspection of Figure 1 shows the 20 categories of interpersonal experience arranged in such a way that they define two major dimensions: positive-negative and active-passive. Thus, the positive section of the positive-negative dimension is defined by the 10 categories A through E and K through O on the right side of Figure 1, while the negative section is defined by the 10 categories F through J and P through T on the left. This arrangement is, of course, arbitrary in the sense that it was based upon the investigator's judgment of the interrelationships among the categories, e.g., the decision to include Category R (distrust, demand, accuse) in the definition of passivity might be challenged. Second, we assumed that these 20 categories of interpersonal experience, regardless of differences among observers or patients to be rated, could be grouped into four clusters to define the two dimensions: active-passive and positive-negative (friendly-hostile). These two dimensions, it will be recalled, are included among the three polarities in terms of which Freud (1957) described the vicissitudes of instinct. In his discussion the two dimensions are used for the description of responses to intrapsychic stimuli (instincts). The studies of Osgood, Suci, and Tannenbaum (1957) and Schaefer (1961) suggest that these are important dimensions of the individual's response to his environment. In short, the work cited above explores the possibility that friendly-hostile and active-passive are the two most important of an invariant system of dimensions which characterize responses to internal and external stimuli.

(F) Dictate, Dominate, Boss
(G) Rebel, Boast, Show Off
(H) Reject, Take Away, Compete
(I) Threaten, Mock, Punish
NEGATIVE ACTIVE RELATIONSHIPS
(J) Disapprove, Attack, Condemn
Direct, Lead, Control Kindly (A)
Teach, Inform, Advise (B)
Give, Interpret, Help (C)
Support, Sympathize, Pity (D)
POSITIVE ACTIVE RELATIONSHIPS
Love, Praise, Approve (E)
(P) Passively Critical, Resentful, Unappreciative
NEGATIVE PASSIVE RELATIONSHIPS
(Q) Complain, Neg. Resist Passively
(R) Distrust, Demand, Accuse
(S) Apologetic, Self-Critical, Retreating
(T) Cowed into Obedience, Obeying a Feared Authority, Submitting
Loves to be praised, Appreciates (K)
Loves to be liked
POSITIVE PASSIVE RELATIONSHIPS
Cooperate, Conciliate, Agree (L)
Trust, Ask Help, Depend (M)
Admire, Ask Advice, Ask Opinion (N)
Conform, Like to do as others do, (O) Obedient

FIG. 1. A system for classifying interpersonal experiences.

[1] This study is part of a major project "Observations and Concepts Used in Clinical Case Description" supported by the National Institute of Mental Health, United States Public Health Service, under Grant M-2224. We are greatly indebted to R. Jarrett, M. B. Smith, and R. Tryon, University of California, for helpful comments and advice on problems of methodology.

The first hypothesis to be tested here is whether this invariant system holds for clinicians with different professional training responding to the same patient and for clinicians with comparable training responding to different patients. If findings support the hypothesis of an invariant system of dimensions, then the second question raised here is whether these dimensions are equally meaningful to all clinical observers and in observations about all patients. In other words, given the fact that the two dimensional system characterizes, in general, the thinking of several groups of clinicians about a patient, what modifications in the use of the system are associated with certain kinds of professional training for the clinician? In what way is the system adapted when professional training is held constant and the patient to be described is varied? In order to spell out the modifications associated with observer or patient differences we would wish to know, first, whether positive-negative and active-passive are two dimensions equally characteristic of clinical thinking regardless of professional training or patient described.

Second, there may still be qualitative differences in the meaning assigned to the positive-negative or the active-passive dimensions. For instance, in the a priori system shown in Figure 1, Categories F (dictate, dominate, boss) and A (direct, lead, control kindly) define one pole of the active-passive dimension, while Categories T (cowed into obedience, obeying a feared authority, submitting) and O (conform, like to do as others do, obedient) define the other. If the meaning of this dimension for a given group of clinicians in the context of a given case is the same as the one in the a priori system, then Categories F and A should represent the greatest possible contrast to Categories T and O when clinicians describe a patient, regardless of whether this patient is thought to be extremely passive or just average and regardless of the degree of variability in judgments about the patient. However, some observers may consider Category I (threaten, mock, punish) and B (teach, inform, advise) a much better contrast to Categories T and O. In short, differences in the quality of meaning assigned to the two dimensions depend upon the relationships perceived among the categories which define them. The magnitude of the correlations which measure these relationships depends upon the degree of systematic variation found in the respnses of a given sample of observers. Such variation in response may be a function of the particular group of observers studied, of the persons to be rated by the observers, or of a combination of these two factors. In those studies in which responses to a variety of persons are rated by the same observers, the relationships found among the categories are a function of both differences in thinking processes of the observer and differences in the persons rated (Levy & Dugan, 1960; Osgood et al., 1957). In the present study we have held the person to be rated constant in order to evaluate the thinking processes of various groups of observers.

In this study 301 Freudian clinicians were asked to describe the same patients in terms of the categories shown in Figure 1 in order to see if their use of the categories was consistent with the a priori system and to determine systematic differences related to professional training and contexts of the clinical material.[2] If the categories are used without special bias by clinicians, one would expect to find that the same dimensions are equally important in their description of the patient and

[2] As described in Chance and Arnold (1960), initially 3,564 clinicians from six major urban areas were contacted. Their names had been obtained from membership directories for the American Psychiatric Association, the Clinical Division of the American Psychological Association, and the Psychiatric Social Work section of the National Association of Social Workers. Willingness to participate was indicated by 1,058 of these clinicians by supplying information about their professional training, experience, and preferred theoretical systems. From this group 542 actually evaluated the clinical material, but only 533 followed instructions closely enough to permit use of their responses in this part of the study. This group of 533 clinicians differed from the sample of 3,564 contacted initially in that psychologists had participated significantly more than members of other disciplines. Of this group 186 psychiatrists, 55 psychologists, and 115 social workers cited Freud as the major influence upon their clinical thinking. In this study only the responses of psychiatrists and social workers were used because the *N* of 55 for psychologists was considered inadequate as stated in Footnote 3.

Rating Procedure for describing the content of the patient's verbalizations: The categories listed below should be marked from 1 (least descriptive) to 20 (most descriptive) of the patient's verbalized preoccupation with certain kinds of interpersonal experiences and relationships. First, select out the category which is most characteristic of the patient's verbalizations about his interpersonal experience and assign it the rank of 20. Select the next most characteristic category and assign it the rank of 19. The third most descriptive category should receive the rank of 18. Next, proceed to assign the rank of 1 to the category which least applies to the patient's verbalization about his interpersonal experience. A rank of 2 should be given to the next least descriptive category. Proceed in this fashion until you have used all the ranks from 1 to 20 taking care not to use the same number twice.

We realize that quantification of clinical material raises many problems and is subject to a number of theoretical objections. Nevertheless we hope that for the purpose of this study you will use this procedure to describe your *first impressions*. Please *limit yourself to 15 minutes* for rating the transcript in this way.

Twenty Categories of Interpersonal Experience in Random Order*	Your ranking from 1, least descriptive, to 20, most descriptive
Complain, Nag, Resist passively	.....
Passively critical, Resentful, Unappreciative	.....
Reject, Take away, Compete	.....
Distrust, Demand, Accuse	.....
Cooperate, Conciliate, Agree	.....
Teach, Inform, Advise	.....
Admire, Ask advice, Ask opinion	.....
Disapprove, Attack, Condemn	.....
Conform, Like to do as others do, Obedient	.....
Loves to be praised, Appreciates, Loves to be liked	.....
Dictate, Dominate, Boss	.....
Support, Sympathize, Pity	.....
Rebel, Boast, Show off	.....
Apologetic, Self-critical, Retreating	.....
Threaten, Mock, Punish	.....
Give, Interpret, Help	.....
Cowed into obedience, Obeying a feared authority, Submitting	.....
Love, Praise, Approve	.....
Direct, Lead, Control kindly	.....
Trust, Ask help, Depend	.....

Note.—Four hundred randomizations were systematically distributed throughout the sample of 533 clinicians in the original report in order to eliminate the effect of order of presentation.

FIG. 2. Rating sheet used by clinicians.

that each of the categories receives the same empirical definition. Such a design aims at exploring communality in the extent and quality of meaning for clinicians with different professional training. The same procedure in each group of clinicians, replicated for several patients, permits exploration of questions about stability of meaning related to differences among patients.

METHOD

As reported in an earlier paper (Chance & Arnold, 1960) psychiatrists and social workers from six major urban areas in the United States had been asked to evaluate the transcripts of two treatment sessions: one analytic hour with a middle-aged business man—a sophisticated patient deeply involved in psychoanalysis, and one psychotherapy hour with an ambulatory, schizophrenic girl—promiscuous, alcoholic, and unable to use the hour for more than a report of her daily mishaps. Both patients had been seen by the same therapist for over one year.

Mimeographed transcripts of the two treatment sessions were sent to each participating clinician together with instructions to evaluate the verbalizations of each patient by means of the instrument shown in Figure 2. It will be noted that the categories which the clinicians were asked to rank in order of their descriptiveness for the verbalizations of each patient are the same as those shown in Figure 1, but were presented to the subjects in different random orders. Half of the subjects were instructed to read and evaluate the therapy transcript first; the other half were asked to read and evaluate

TABLE 1

BREAKDOWN OF THE SAMPLE

Sample	Number describing	
	Analytic patient	Therapy patient
Analytically trained psychiatrists	54	51
Psychiatrists without analytic training	37	44
Social workers	62	53

the analytic transcript first. This attempt at counterbalancing the effects of order of case presentation was not successful.[3] Therefore, only responses to the first case read were used in this investigation. Table 1 shows the breakdown of the sample of Freudian psychiatrists and social workers who returned the materials in usable condition. One hundred and fifty-three evaluated the analytic transcript, and 148 the therapy transcript. Within each of these major divisions of the sample the responses of three groups were examined separately: psychiatrists with formal analytic training, psychiatrists without such training, and social workers. We were interested, first, in determining whether given the same patient clinicians with varying professional experience would agree in the meaning attributed to the 20 categories. The second problem was to determine whether holding professional training constant clinicians assessing two different cases would attribute the same meaning to the 20 categories of interpersonal relationships. Technical problems in the investigation of these questions can be considered in two steps: first, the choice of a technique for describing the relationships among the ranks obtained from a given group for the 20 categories in terms of underlying dimensions, and, second, the choice of the most logical procedure for comparing the responses of six groups of clinicians.

Factor analysis was the technique used for analyzing the relationships obtained among rankings for the 20 categories by each group of clinicians in terms of the dimensions underlying these rankings. The first step was to factor the covariance matrices generated from the six sets of category rankings. This was done using the centroid method after correcting for the effects of ipsative scores according to Tucker (1956). The factor loadings thus obtained were then standardized by dividing by the standard deviations of the associated variables in order to obtain results equivalent to a factor analysis of correlation coefficients. In line with the first hypothesis concerning the importance of the two dimensions, it was decided that unless the first two factors from a given analysis could account for at least 25% of the variance present, we were not interested in proceeding further. In addition, since the amount of variance left to be extracted was, in general, minimal after removing the first two centroid factors, it was decided to extract only these two and a third to assist in matching them to the criterion.[4]

In order to clarify the procedures used, the model and the hypothesis are presented in some detail. The main hypothesis is that for all English speaking clinicians there exists an invariant semantic structure associated with the 20 categories presented in Figure 1. That is to say, that within the limits imposed by errors of sampling and measurement, there exists an orthogonal rotation of the factor matrix of one group of clinicians to a position where it will correlate perfectly with the factor matrix of another group of clinicians. Note that this says nothing about the relative amount of variance associated with the various factors. The theoretical correlations are not affected by this. The test of the hypothesis is to rotate one factor matrix to maximum agreement with another and determine whether the correlation between the corresponding factor loadings departs materially from 1.00. A degree of discretion is required here since departure from 1.00 will to some extent be a function of the reliability of the loadings, and loadings on factors that extract smaller proportions of variance may be expected to be less reliable than those on factors that extract larger amounts of variance.

For comparisons of factorial results from any one group with those from another, reference axes for the factor matrix of that group were rotated to a least squares solution [5] of maximum agreement with results for the other group. The comparisons reported here involved the rotations listed in Table 2. In these comparisons the responses of analytically trained psychiatrists to the analytic case were selected as the criterion performance. It was assumed that this particular group of clinicians had the longest and most intensive training and that their responses would be most representative of Freudian clinical thinking. Responses to the analytic transcript were selected as the criterion performance since the session with that patient represented a more classically Freudian instance of therapy than the session with the psychotherapy patient. As shown in Table 2 studies of communality of meaning among observers with varying professional training for the analytic case involved comparison of the responses of psychiatrists without analytic training and of those of social workers with the criterion performance. For studies of communality of meaning in the description of the therapy patient the responses of psychoanalytically trained psychiatrists for that patient were first rotated toward the criterion performance in order to obtain the most "typically" Freudian organization of the data. The responses to the therapy

[3] It was demonstrated that order of case presentation showed marked effects upon the clinicians' judgment about each patient's concern with negative-passive relationships. If a patient was assessed first, he was thought to be more preoccupied with resentment, complaint, and distrust than if he was assessed second.

[4] The factor loadings were stabilized by two separate factorings of the covariance matrices using loadings obtained in the first factoring of categories to estimate the communalities for the second. The differences between the first and second estimates were so slight that the results of the second were accepted, and the procedure was carried no further.

[5] The reference axes were rotated to minimize the squared distances between comparable vectors (variables). The solution requires maximizing the inner products (covariances) of comparable factor vectors. Orthogonal solutions were used. Since we had no hypothesis concerning the relationship between the two factors, orthogonality appeared the simplest and most parsimonious assumption.

TABLE 2

MATRIX ROTATIONS FOR COMPARING RESULTS OF FACTORIAL STUDIES

Observers	Analytic patient	Therapy patient
Matrix rotations for studies of communality of meaning among observers with varying professional training		
Analytically trained psychiatrists	Matrix 1 Criterion (No rotation)	Matrix 2 Rotated to criterion Matrix 1 (Criterion for other therapy-patient matrices after rotation)
Psychiatrists without analytic training	Matrix 3 Rotated to criterion Matrix 1	Matrix 4 Rotated to criterion Matrix 2
Social workers	Matrix 5 Rotated to criterion Matrix 1	Matrix 6 Rotated to criterion Matrix 2
Matrix rotations for studies of stability of meaning in the context of different patients		
Analytically trained psychiatrists	Matrix 1 Criterion (No rotation)	Matrix 2 Rotated to criterion Matrix 1
Psychiatrists without analytic training	Matrix 3 Rotated to criterion Matrix 1 (Criterion after rotation for therapy-patient matrix of similar group)	Matrix 4 Rotated to criterion Matrix 3
Social workers	Matrix 5 Rotated to criterion Matrix 1 (Criterion after rotation for therapy-patient matrix of similar group)	Matrix 6 Rotated to criterion Matrix 5

transcript of psychiatrists without analytic training and of social workers were then compared with those of the analytically trained psychiatrists.

In order to determine the extent to which this sample of Freudian clinicians attributed the same meaning to the 20 categories in the context of describing the analytic and the therapy transcripts, responses for the analytic patient by psychiatrists without analytic training and by social workers were first rotated toward the criterion performance of analytically trained psychiatrists. Having thus obtained the most "typically" Freudian arrangement of the categories, the responses of each of the three groups of clinicians for the analytic patient were then compared with those of the corresponding groups for the therapy patient.

In all comparisons, regardless of whether they concerned communality of meaning among observers with different professional training or stability of meaning in the description of different patients, the phenomena of interest can be described in quantitative terms[6] as follows:

1. The importance of the two-dimensional system for the thinking of each group of clinicians can be measured in terms of the average proportion of total variance explained by the positive-negative and the active-passive factors.

2. Whereas the importance of the system as a whole may remain unaffected by variations in the professional training of observers or by variations in clinical material described, the relative importance of the two factors may be affected by such variations. Variations in the ratio of the average propor-

[6] Statistical tests of differences between variance-covariance matrices were not undertaken because the basic data were ranks, thus rendering the matrices only approximations of what might have been obtained with scores with less severe restrictions on them.

tion of the total variance explained by the two factors were, therefore, examined for evidence of the effects of different professional training or different clinical material.

3. Similarity in the meaning of the two dimensions for clinicians with different professional training or for the description of different patients can be measured as follows: The relationship of each category to an underlying factor such as the positive-negative dimension is described by its loading on that factor. Factor loadings obtained for a given matrix for each of the 20 categories on the positive-negative dimension can be correlated with those obtained from the criterion matrix. Product-moment correlations were used.

4. While the above correlations may serve as measures of similarity in the over-all meanings assigned to the dimensions, this method fails to pinpoint the kind of differences in the meaning of the dimensions which might occur among different observers or in the context of different patients. In order to explore such qualitative shifts in meaning, the dimensions were arbitrarily considered to be empirically defined for each group of observers by the five categories at each pole (i.e., those with the highest positive and the highest negative algebraic factor loadings). For example, for the criterion group of psychoanalytically trained psychiatrists, the five most active and the five most passive categories define the dimension of activity-passivity in the context of a given case. Comparison of this definition of activity-passivity with that produced by a group with similar training in the course of describing the therapy patient reveals shifts in meaning related to shifting clinical context. Comparison of the criterion group's definition of the dimension in the context of the analytic case with definition of the dimension in the same context produced by another group of observers reveals shifts in meaning related to professional training.

TABLE 3

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE ANALYTICALLY TRAINED PSYCHIATRISTS TO THE ANALYTIC TRANSCRIPT

Category	Factor			h^2
	I	II	III	
A (Lead)	.31	.57	.14	.44
B (Teach)	.41	.33	.31	.37
C (Give)	.57	.48	−.22	.60
D (Support)	.44	.42	.17	.40
E (Love)	.46	.21	−.12	.27
F (Boss)	−.23	.49	−.25	.36
G (Rebel)	−.40	.51	.31	.52
H (Reject)	−.39	.30	.33	.35
I (Punish)	−.52	.23	−.20	.36
J (Attack)	−.66	.14	−.43	.64
K (Appreciates)	.28	−.19	−.26	.18
L (Cooperate)	.52	−.36	.27	.47
M (Trust)	.51	−.29	−.19	.38
N (Admire)	.48	−.32	−.25	.40
O (Conform)	.36	−.48	.24	.42
P (Resentful)	−.58	−.26	.43	.59
Q (Complain)	−.42	−.29	.19	.30
R (Accuse)	−.59	−.06	−.25	.41
S (Retreating)	−.13	−.59	−.22	.41
T (Submitting)	−.37	−.54	.16	.45
Sum of squares	4.05	2.93	1.35	
Average proportion of total variance	.18	.12	.06	
Average proportion of common variance	.50	.33	.17	

RESULTS AND DISCUSSION

Findings[7] Concerning Communality of Meaning among Observers with Different Professional Training

Descriptions of the analytic patient's verbalizations: The results of factoring the responses of the criterion group of analytically trained psychiatrists, of psychiatrists without analytic training, and of social workers are shown in Tables 3, 4, and 5, respectively. It will be noted that in each of these three factor analyses the first two factors extracted were the positive-negative and active-passive dimensions postulated in the system shown in Figure 1. The importance of the two-dimensional system as measured by the average proportion of the total variance explained by the first two factors ranged from .30 for the criterion group to .32 for the group of social workers, failing to differentiate among the three groups.

The relative importance assigned to the two factors did differentiate among clinicians. The criterion group of analytically trained psychiatrists assigned greater importance to the positive-negative than to the active-passive dimension. The social workers assigned equal importance to the two dimensions. In contrast to this psychiatrists without analytic training assigned greater importance to the active-passive dimension than to the positive-negative dimension. The emphasis on the positive-negative relative to the active-pas-

[7] Statistical data supplementary to this paper have been deposited with the American Documentation Institute. Order Document No. 7092 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $2.00 for microfilm or $3.75 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

TABLE 4

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE PSYCHIATRISTS WITHOUT ANALYTIC TRAINING TO THE ANALYTIC TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.02	.39	−.01	.15
B (Teach)	.10	.74	.17	.59
C (Give)	.39	.49	−.29	.48
D (Support)	.25	.06	.11	.08
E (Love)	.23	.16	−.43	.26
F (Boss)	−.49	.59	.03	.59
G (Rebel)	−.16	.37	.14	.18
H (Reject)	−.33	.12	.15	.15
I (Punish)	−.06	.52	−.61	.65
J (Attack)	−.19	.36	−.40	.33
K (Appreciates)	.38	−.33	.46	.46
L (Cooperate)	.60	−.06	−.11	.38
M (Trust)	.37	−.34	.01	.25
N (Admire)	.52	−.24	.15	.35
O (Conform)	−.08	−.48	−.20	.28
P (Resentful)	−.40	−.18	.67	.64
Q (Complain)	−.38	−.31	.46	.45
R (Accuse)	−.72	.12	.13	.55
S (Retreating)	−.14	−.77	.00	.61
T (Submitting)	−.05	−.87	−.18	.79
Sum of squares	2.46	3.88	1.88	
Average proportion of total variance	.12	.19	.09	
Average proportion of common variance	.30	.47	.23	

Transformation matrix

.264282	−.909282	.321498
.961968	.224658	−.155415
.069124	.350344	.934070

[a] Rotated to matrix of analytically trained psychiatrists for analytic transcript.

sive factor can be expressed for the three groups by the following ratios: 1.5, 1.0, and .63. Psychiatrists without analytic training presented an almost exact reversal of emphasis compared with psychiatrists with formal analytic training.

Shifts in the over-all meaning of the two dimensions were affected by differences in the professional training of observers as follows: Correlations between the responses of the criterion group and the psychiatrists without analytic training and social workers were .83 and .90, respectively, for the factor loadings of categories on the positive-negative dimension, and .67 and .92 for those on the active-passive dimension. Although differences among these correlations cannot be evaluated statistically, the data suggest that the social workers resembled the analytically trained psychiatrists more in their use of these dimensions than psychiatrists without psychoanalytic training.

Empirical definitions of each dimension by each group of observers are illustrated in Figures 3 and 4. Figure 3 shows the five categories considered most positive and the five considered most negative by each group of observers. Inspection of the figure shows that all three groups of clinicians agree in defining this dimension in terms of the polarity: trust, ask help, depend (Category M)-distrust, accuse, demand (Category R).

TABLE 5

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE SOCIAL WORKERS TO THE ANALYTIC TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.27	.47	.30	.38
B (Teach)	.26	.38	.33	.32
C (Give)	.24	−.02	.19	.09
D (Support)	.43	.31	.21	.33
E (Love)	.54	.23	−.47	.57
F (Boss)	−.24	.79	−.24	.74
G (Rebel)	−.33	.56	.14	.44
H (Reject)	−.52	.26	−.21	.38
I (Punish)	−.49	.40	−.32	.50
J (Attack)	−.81	.07	−.02	.66
K (Appreciates)	.49	−.12	−.09	.26
L (Cooperate)	.21	−.41	.34	.33
M (Trust)	.28	−.41	−.23	.30
N (Admire)	.27	−.46	−.16	.31
O (Conform)	.48	−.40	.03	.39
P (Resentful)	−.27	−.36	.55	.51
Q (Complain)	−.21	−.40	.41	.37
R (Accuse)	−.64	−.05	−.38	.56
S (Retreating)	.02	−.64	−.03	.41
T (Submitting)	.01	−.36	−.16	.16
Sum of squares	3.20	3.24	1.57	
Average proportion of total variance	.16	.16	.08	
Average proportion of common variance	.40	.40	.20	

Transformation matrix

.647449	−.750007	.135273
.748315	.659254	.073546
−.144339	.053608	.988075

[a] Rotated to matrix of analytically trained psychiatrists for analytic transcript.

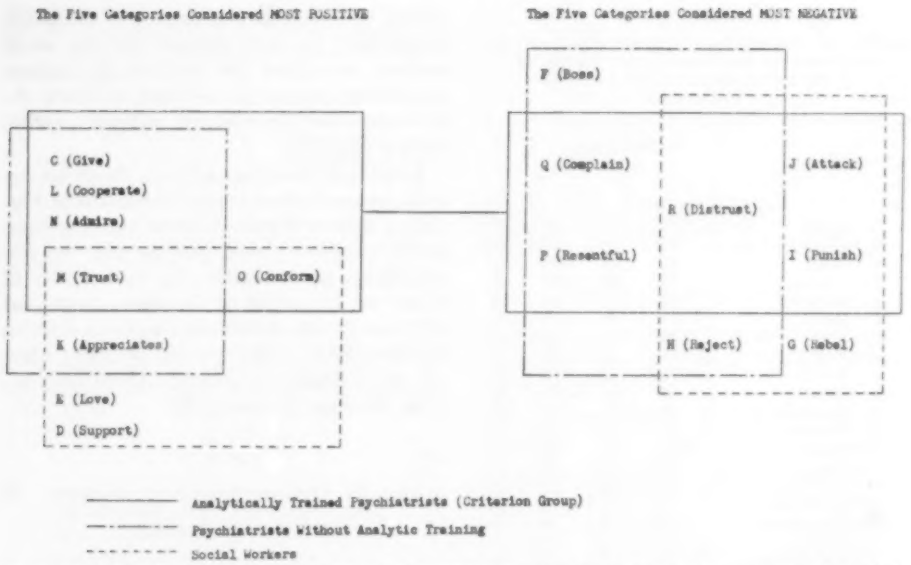

FIG. 3. Comparisons of empirical definitions of the positive-negative dimension in the context of the analytic patient.

Figure 4 shows empirical definitions for the active-passive dimension. Again it will be noted that of the 10 categories which constitute the criterion group's definition of this dimension all but 1 were also used by either one or both of the other two groups of cli-

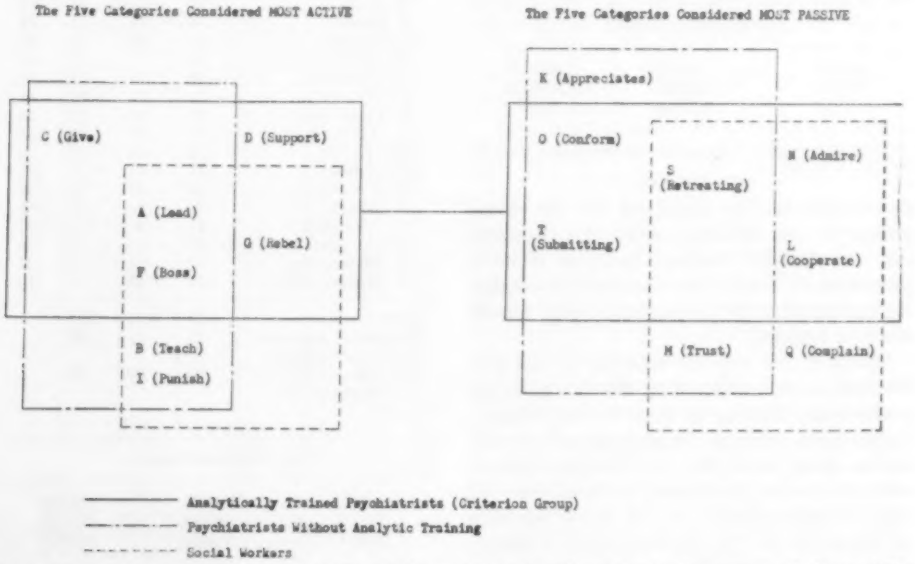

FIG. 4. Comparisons of empirical definitions of the active-passive dimension in the context of the analytic patient.

TABLE 6

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE ANALYTICALLY TRAINED PSYCHIATRISTS TO THE THERAPY TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.56	.31	.21	.45
B (Teach)	.28	.54	.28	.45
C (Give)	.63	.40	.11	.57
D (Support)	.49	.30	−.21	.37
E (Love)	.55	.12	−.31	.41
F (Boss)	−.43	.19	.17	.25
G (Rebel)	−.42	.54	−.22	.52
H (Reject)	−.57	−.02	.30	.42
I (Punish)	−.49	.27	−.30	.40
J (Attack)	−.63	−.02	.05	.40
K (Appreciates)	.32	−.11	−.48	.34
L (Cooperate)	.36	−.28	.13	.22
M (Trust)	.47	−.26	.25	.35
N (Admire)	.35	−.29	.15	.23
O (Conform)	.32	−.59	−.11	.46
P (Resentful)	−.49	−.08	.19	.28
Q (Complain)	−.67	−.06	.16	.48
R (Accuse)	−.72	−.03	.19	.56
S (Retreating)	−.02	−.62	−.18	.42
T (Submitting)	.11	−.30	−.23	.16
Sum of squares	4.55	2.13	1.06	
Average proportion of total variance	.23	.11	.05	
Average proportion of common variance	.59	.27	.14	

Transformation matrix

.976077	−.190091	−.105533
.163094	.961131	−.222772
.143778	.200230	.969142

[a] Rotated to matrix of analytically trained psychiatrists for analytic transcript.

nicians. The use of Category D (support, sympathize, pity) as part of their definition of activity distinguished analytically trained psychiatrists from the other two groups. Both social workers and psychiatrists without analytic training differed from the criterion group in using Categories B (teach, inform, advise) and I (threaten, mock, punish) in their definition of activity. Both also differed from the criterion group in that they included Category M (trust, ask help, depend) in their definition of passivity. In summary, empirical definitions of the active-passive dimension appeared in the context of the analytic case to differentiate more sharply between analytically trained and other clinicians than the definitions of the positive-negative dimension.

Descriptions of the therapy patient's verbalizations: The results of factoring responses of the criterion group of analytically trained psychiatrists, of psychiatrists without analytic training, and of social workers are shown in Tables 6, 7, and 8. Findings for responses to the therapy patient were similar to those for the analytic patient in that the first two factors extracted in each of three factor analyses were the positive-negative and active-passive dimensions postulated in the a priori system. The importance of the two-dimensional system as measured by the average proportion of total variance explained by the first two factors represented a wider range: .34

TABLE 7

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE PSYCHIATRISTS WITHOUT ANALYTIC TRAINING TO THE THERAPY TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.15	.47	.14	.26
B (Teach)	.54	.47	.04	.51
C (Give)	.72	−.04	.22	.57
D (Support)	.59	−.29	.33	.54
E (Love)	.84	.05	−.03	.71
F (Boss)	−.64	.40	.49	.81
G (Rebel)	−.49	.47	−.07	.47
H (Reject)	−.75	.28	−.05	.64
I (Punish)	−.64	.10	−.26	.49
J (Attack)	−.81	−.03	.15	.68
K (Appreciates)	.37	−.20	−.62	.56
L (Cooperate)	.41	−.25	−.23	.28
M (Trust)	.78	−.05	−.06	.61
N (Admire)	.67	.22	−.08	.50
O (Conform)	.39	−.28	−.30	.32
P (Resentful)	−.69	−.12	.33	.60
Q (Complain)	−.53	−.30	.38	.52
R (Accuse)	−.73	.26	−.18	.63
S (Retreating)	.19	−.42	.02	.21
T (Submitting)	.06	−.32	−.03	.11
Sum of squares	7.01	1.68	1.34	
Average proportion of total variance	.35	.08	.07	
Average proportion of common variance	.70	.17	.13	

Transformation matrix

.938945	−.298356	−.171365
.342764	.767823	.541258
−.029910	−.566949	.823210

[a] Rotated to rotated matrix of analytically trained psychiatrists for therapy transcript.

TABLE 8

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE SOCIAL WORKERS TO THE THERAPY TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.41	.13	.06	.19
B (Teach)	.20	.12	−.30	.14
C (Give)	.67	.05	.15	.47
D (Support)	.64	−.17	.25	.50
E (Love)	.64	.00	−.27	.48
F (Boss)	−.73	.04	−.02	.53
G (Rebel)	−.52	.18	.15	.33
H (Reject)	−.61	.18	−.11	.42
I (Punish)	−.64	.35	−.25	.59
J (Attack)	−.67	.35	.12	.59
K (Appreciates)	.43	.35	−.23	.36
L (Cooperate)	.41	−.40	−.10	.34
M (Trust)	.71	.21	.11	.56
N (Admire)	.64	.01	.24	.47
O (Conform)	.31	−.44	−.39	.44
P (Resentful)	−.55	−.10	−.09	.32
Q (Complain)	−.61	−.27	.30	.54
R (Accuse)	−.63	.22	.28	.52
S (Retreating)	.28	−.40	.29	.32
T (Submitting)	−.10	−.20	−.17	.08
Sum of squares	6.03	1.23	.94	
Average proportion of total variance	.30	.06	.05	
Average proportion of common variance	.74	.15	.11	

Transformation matrix

.990466	−.128751	−.048996
.130223	.991086	.028128
.044938	−.034277	.998403

[a] Rotated to rotated matrix of analytically trained psychiatrists for therapy transcript.

for the criterion group, .36 for social workers, and .43 for psychiatrists without analytic training. In their use of the two-dimensional system social workers resembled analytically trained psychiatrists more than did psychiatrists without analytic training.

In the description of the therapy patient by each group of clinicians the emphasis on the positive-negative dimension was greater than that on the active-passive dimension. However, there was marked contrast between analytically trained psychiatrists and the other two groups in the relative emphasis placed upon this dimension. The ratio for the criterion group was 2.1, whereas it was 4.4 and 5.0, respectively, for psychiatrists without analytic training and for social workers.

The over-all meaning of the positive-negative dimension was not affected by differences in the professional training of clinicians. Correlations of .94 and .96 were obtained between the responses of the criterion group and psychiatrists without analytic training on the one hand and social workers on the other. In contrast to this, the over-all meaning of the active-passive dimension was considerably affected by differences in the professional training of observers. The correlation between the responses of psychiatrists with and without analytic training was .65; that for the responses of social workers with the criterion group was .53.

An examination of the empirical definitions of each dimension by the three groups of clinicians reveals the areas of agreement as well as differences in the quality of meaning attributed to the two dimensions in the context of describing the therapy patient's verbalizations. Figure 5 shows that of the 10 categories used by the criterion group in their definition of the positive-negative dimension all but 1 were used by either or both of the other two groups. As might be expected in view of the high correlations among factor loadings all three groups agreed in defining the positive pole of the dimension in terms of 4 of the 5 categories used for definition of that pole. Agreement concerning the definition of the negative pole was less striking. Empirical definitions of the active-passive dimension are shown in Figure 6. Of the 10 categories used by the criterion group in defining this dimension 8 were also used by either or both of the other two groups. Clinicians with different professional training showed marked differences in defining this dimension. Social workers used an entirely different set of 5 categories from those employed by the other two groups in their definition of activity, but in defining passivity they agreed with the criterion group in the choice of 4 out of 5 categories. While psychiatrists without analytic training agreed with their analytically trained colleagues in the choice of 3 of the 5 categories defining each pole, there was one instance of striking disagreement between the two groups. While analytically trained psychiatrists included Category D (support, sympathize,

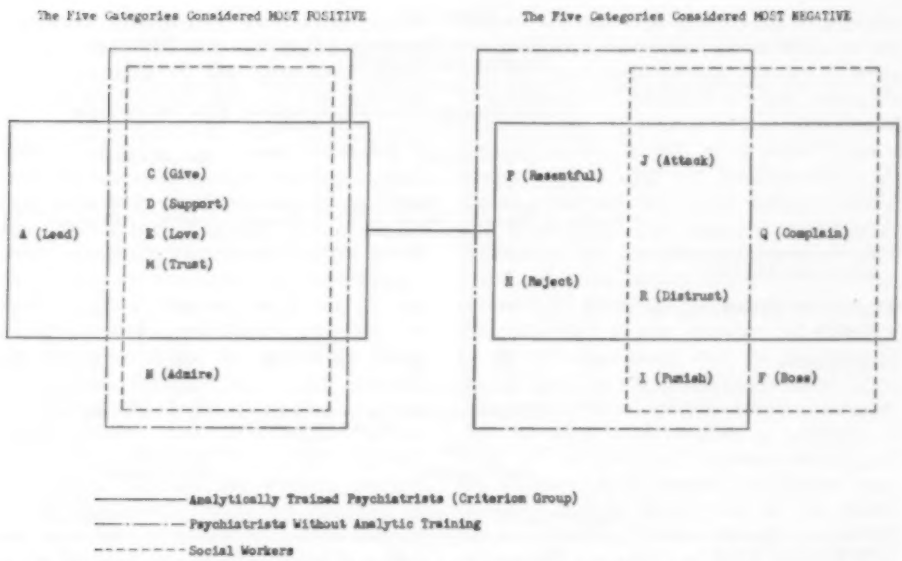

FIG. 5. Comparisons of empirical definitions of the positive-negative dimension in the context of the therapy patient.

pity) in their definition of activity, psychiatrists without analytic training included the same category in their definition of passivity.

Table 9 summarizes quantitative findings concerning communality of meaning among clinicians with different professional training.

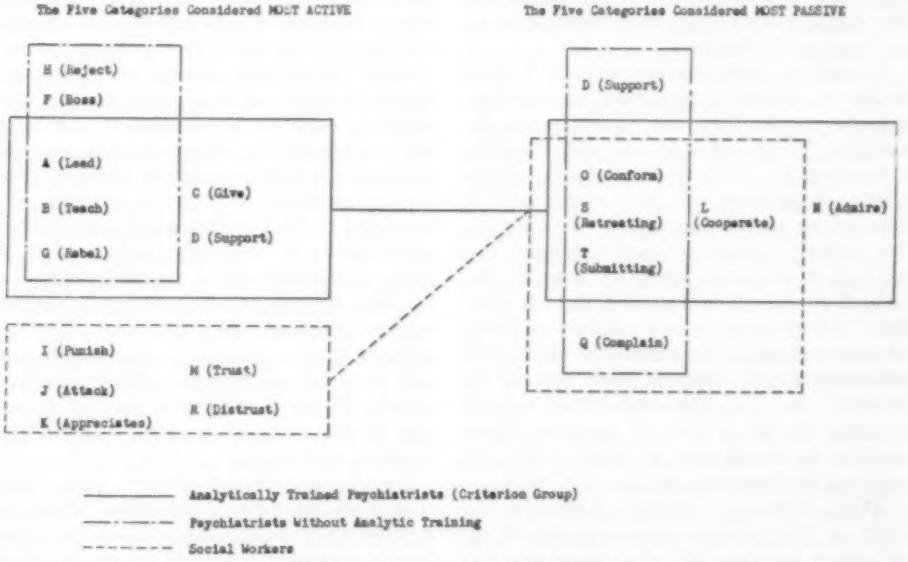

FIG. 6. Comparisons of empirical definitions of the active-passive dimension in the context of the therapy patient.

TABLE 9

SUMMARY OF FACTORIAL STUDIES COMPARING RESPONSES OF CLINICIANS WITH DIFFERENT PROFESSIONAL TRAINING

	Average proportion of total variance explained by the first two factors							
	Analytic patient				Therapy patient			
	Positive-Negative	Active-Passive	Both factors	Ratio[a]	Positive-Negative	Active-Passive	Both factors	Ratio[a]
Criterion group (Analytically trained psychiatrists)	.18	.12	.30	1.5	.23	.11	.34	2.1
Psychiatrists without analytic training	.12	.19	.31	.63	.35	.08	.43	4.4
Social workers	.16	.16	.32	1.0	.30	.06	.36	5.0

	Communality of meaning for each factor measured by correlation between factor loadings			
	Analytic patient		Therapy patient	
	Positive-Negative	Active-Passive	Positive-Negative	Active-Passive
Criterion group versus Psychiatrists without analytic training	.83	.67	.94	.65
Criterion group versus Social workers	.90	.92	.96	.53

[a] Ratio: Positive-Negative/Active-Passive.

The table shows, first, that the postulated two factor system was indeed meaningful to each group of clinicians.

Second, the table shows a tendency in four of the six groups to emphasize the positive-negative rather than the active-passive dimension. This trend is in line with Osgood's (Osgood et al., 1957) findings concerning the relative importance of three dimensions in descriptions of a large variety of concepts: "evaluation," "potency," and "activity." Regardless of the scales used, the concepts described, or the subjects studied, Osgood finds that "differentiation among concepts in terms of their evaluation is about twice as fine as differentiation in terms of their potency or activity" (p. 73). This same trend can be observed in 16 of the 19 separate factor analyses for descriptions of different concepts reported by Osgood et al. (pp. 181–185).

Third, contrary to Osgood's (Osgood et al., 1957, p. 223) findings concerning the effects of subject selection, the table shows that clinicians differed in the relative emphasis placed upon the two dimensions. These differences were a function of both professional training and the case described. In the context of describing the analytic patient social workers performed more as analytically trained psychiatrists, while in the context of describing the therapy patient they performed more as psychiatrists without analytic training. This apparent tendency to a shift in emphasis may be related to the fact that social workers have opportunities to work with and to learn to adapt to the methods of case description expounded by psychiatrists with and without analytic training. The fact that they responded more as analysts in the analytic case and more as psychiatrists without analytic training in the therapy case may be the result of this capacity to adapt to their professional environment.

Fourth, examination of the table shows that the active-passive dimension, which accounted for a smaller proportion of the variation in responses in four of the six groups of clinicians, is also the dimension which tends

to be used differently by analytically trained and untrained clinicians. Agreement upon the over-all meaning of the dimension ranged from .53 to .92 with three of the four correlations obtained at .67 or below.

Empirical definitions of each dimension in terms of the five categories assigned to each of the poles were examined in order to determine whether the six groups differed in the quality of meaning attributed to the dimensions. It was not surprising that for this sample of Freudian clinicians the groups of analytically trained psychiatrists provided the most generally agreed on definitions. However, in their definition of activity for each case analytically trained psychiatrists consistently differed from the other two groups by including Category D (support, sympathize, pity). Psychiatrists without analytic training consistently used this category as more indicative of passivity and in describing the therapy case even included it among the five most passive attributes which constitute their empirical definition of passivity. A speculative explanation of this difference may be found in examining the definitions of activity which distinguished analytically trained and untrained clinicians. In the context of the analytic case both psychiatrists without analytic training and social workers differed from the criterion group in their inclusion of Categories B (teach, inform, advise) and I (threaten, mock, punish) in definitions of activity. It is possible that Category D (support, sympathize, pity) was contrasted with Categories B and I in the context of the analytic material because clinicians without analytic training may see the psychoanalyst primarily in a passive role.

Definitions of the active-passive dimension in the context of the therapy material presented a more complex picture than in the context of the analytic material. The three groups of clinicians appeared to differ more in their definitions of the active than of the passive pole for this dimension. Whereas the definitions of activity obtained from psychiatrists with and without analytic training contained some of the same categories, the social workers provided an entirely different definition from that produced by the other two groups. Since the therapy patient was a woman and since the two groups of psychiatrists differed from the social workers not only in professional training but also in terms of a marked predominance of males, one questions whether differences in the definition of activity in the context of material from a woman patient may not be responsible for the differences in the social workers' definition of activity. Definitions of passivity in the context of the therapy material show more agreement among clinicians. However, exclusion of Category Q (complain, nag, resist passively) in the definition of passivity by the criterion group and the inclusion of this category in the definitions obtained from clinicians without analytic training suggests that the latter tend to think of passivity in terms of the manifestly passive meaning of the category, while analytically trained psychiatrists may be more aware of the potential for controlling others through complaints and passive resistance.

In summary, it was possible to demonstrate that using the categories shown in Figure 1 clinical case description, like descriptions of maternal and child behavior (Schaefer, 1961) or descriptions of concepts in general (Osgood et al., 1957), is characterized by two dimensions labeled here positive-negative and active-passive. Communality of meaning for Freudian clinicians with different professional training was greater for the positive-negative than for the active-passive dimension. Differences in meaning were related partly to differences in professional training and partly to the nature of the clinical material. In order to explore the effects of shifts in clinical context upon the stability of meaning of the two dimensions, the responses of the six groups were further analyzed as reported below.

Findings Concerning Stability of Meaning in the Context of the Analytic and the Therapy Patients' Verbalizations

As outlined in Table 2 studies of stability of meaning in the evaluation of the two patients by the six groups of clinicians involved the following procedures: As a first step, the factor matrices for responses to the analytic data obtained from psychiatrists without analytic training and from social workers were

TABLE 10

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE PSYCHIATRISTS WITHOUT ANALYTIC TRAINING TO THE THERAPY TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.20	.48	.01	.27
B (Teach)	.57	.41	−.09	.50
C (Give)	.72	−.04	.21	.56
D (Support)	.58	−.25	.39	.55
E (Love)	.84	−.03	−.06	.71
F (Boss)	−.59	.56	.39	.81
G (Rebel)	−.46	.48	−.18	.47
H (Reject)	−.73	.32	−.11	.65
I (Punish)	−.64	.09	−.26	.49
J (Attack)	−.81	.08	.16	.69
K (Appreciates)	.33	−.38	−.56	.57
L (Cooperate)	.38	−.33	−.16	.28
M (Trust)	.78	−.13	−.05	.63
N (Admire)	.69	.14	−.14	.52
O (Conform)	.35	−.38	−.23	.32
P (Resentful)	−.69	.03	.36	.61
Q (Complain)	−.54	−.15	.45	.52
R (Accuse)	−.71	.27	−.23	.63
S (Retreating)	.15	−.42	.12	.21
T (Submitting)	.04	−.32	.06	.11
Sum of squares	6.84	1.93	1.32	
Average proportion of total variance	.34	.10	.07	
Average proportion of common variance	.67	.20	.14	

Transformation matrix

.904651	−.412585	−.106680
.423888	.845414	.324951
−.043882	−.339189	.939695

[a] Rotated to rotated matrix of psychiatrists without analytic training for analytic transcript.

equated as nearly as possible with the factor matrix for responses from the criterion group of analytically trained psychiatrists.[8] These three factor matrices were then used as criterion matrices in order to evaluate shifts in meaning of the two dimensions associated with the assessment of the therapy patient. Factor matrices for responses to the therapy data were, therefore, rotated to these three matrices as shown in Tables 6, 10, and 11.

Findings concerning stability of meaning attributed to the two dimensions in the context of the analytic and the therapy data are summarized in Table 12. Reference to this table shows, first, that regardless of differences in professional training, the two-dimensional system accounted on the average for a larger proportion of the total variation in response in the context of the therapy data. This suggests the hypothesis that the system of classifying interpersonal experiences may be more descriptive of therapy data than of analytic data.

Second, the table shows that regardless of differences in professional training, differences in clinical context involved shifts in the relative emphasis placed upon the two dimensions. In the context of the therapy data the positive-negative dimension explained more and the active-passive dimension less of the total average variation in response

TABLE 11

MATRIX OF STANDARDIZED FACTOR LOADINGS FOR RESPONSES BY THE SOCIAL WORKERS TO THE THERAPY TRANSCRIPT

Category	Rotated factors[a]			h^2
	I	II	III	
A (Lead)	.42	.09	.09	.19
B (Teach)	.25	.07	−.28	.15
C (Give)	.65	.00	.22	.47
D (Support)	.58	−.21	.33	.49
E (Love)	.67	−.09	−.19	.49
F (Boss)	−.72	.11	−.11	.54
G (Rebel)	−.51	.24	.07	.32
H (Reject)	−.57	.23	−.19	.41
I (Punish)	−.57	.39	−.35	.60
J (Attack)	−.63	.42	.01	.57
K (Appreciates)	.49	.29	−.21	.37
L (Cooperate)	.37	−.45	−.02	.34
M (Trust)	.71	.15	.17	.56
N (Admire)	.61	−.03	.31	.47
O (Conform)	.30	−.51	.31	.45
P (Resentful)	−.55	−.05	−.14	.32
Q (Complain)	−.66	−.18	.25	.53
R (Accuse)	−.62	.30	.18	.51
S (Retreating)	.20	−.39	.35	.31
T (Submitting)	−.10	−.21	−.16	.08
Sum of squares	5.79	1.41	.98	
Average proportion of total variance	.29	.07	.05	
Average proportion of common variance	.71	.17	.12	

Transformation matrix

.970441	−.228096	.078840
.232063	.971643	−.045364
−.066257	.062319	.995854

[a] Rotated to rotated matrix of social workers for analytic transcript.

[8] Results of these rotations are shown in Tables 3, 4, and 5.

TABLE 12

SUMMARY OF FACTORIAL STUDIES COMPARING RESPONSES OF CLINICIANS IN THE DESCRIPTION OF TWO PATIENTS

	Average proportion of total variance explained by the first two factors							
	Analytic patient				Therapy patient			
	Positive-Negative	Active-Passive	Both factors	Ratio[a]	Positive-Negative	Active-Passive	Both factors	Ratio[a]
Criterion group (Analytically trained psychiatrists)	.18	12	.30	1.5	.23	.11	.34	2.1
Psychiatrists without analytic training	.12	.19	.31	.6	.34	.10	.44	3.4
Social workers	.16	.16	.32	1.0	.29	.07	.36	4.1

	Stability of meaning for each factor measured by correlation between factor loadings	
	Positive-Negative	Active-Passive
Analytically trained psychiatrists.	.94	.91
Psychiatrists without analytic training	.83	.76
Social workers	.90	.54

[a] Ratio: Positive-Negative/Active-Passive.

than in the context of the analytic data. These shifts in emphasis are most plausibly explained by reference to the fact that the analytic patient was a man and the therapy patient a woman. In the assessment of female patients clinicians may attribute greater importance to the positive-negative dimension and less importance to the active-passive dimension.

Third, the table shows that regardless of differences in professional training the overall meaning attributed to the positive-negative dimension was less dependent upon differences in clinical context than that attributed to the active-passive dimension.

Fourth, the table suggests that differences in professional training may be associated with differences in the stability of meaning attributed to the two dimensions. Whereas correlations between factor loadings for responses to the analytic and the therapy data were .94 and .91 for the criterion group of analytically trained psychiatrists, corresponding correlations obtained for data from psychiatrists without analytic training and from social workers were lower.

The extent to which the quality of meaning attributed to each dimension was affected by differences in clinical context can be evaluated by examining the empirical definitions obtained from each of the six groups of clinicians. The five categories assigned to each pole of the positive-negative dimension by the six groups are shown in Figure 7. Reference to the figure shows, first, that regardless of differences in professional training and differences in clinical context Category M (trust, ask help, depend) defined the positive pole and Category R (distrust, demand, accuse) defined the negative pole of the dimension. Thus, the meaning of the positive-negative dimension was defined in terms of the patient's readiness to relate to people, presumably particularly to therapists.

Empirical definitions of the active-passive dimension are shown in Figure 8. The figure shows that only the passive pole of this dimension could be defined regardless of differences in professional training or clinical context and that passivity was designated by a single category, S (retreating, apologetic, self-critical). Groups of clinicians differed in the extent to which stability of the quality of meaning attributed to this dimension was

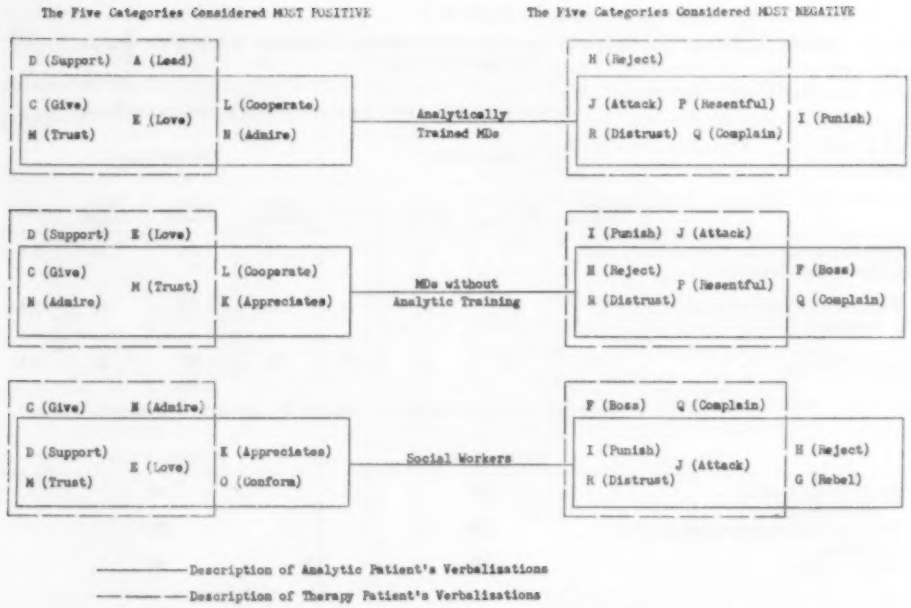
The Five Categories Considered MOST POSITIVE
The Five Categories Considered MOST NEGATIVE
D (Support) A (Lead)
C (Give) E (Love) L (Cooperate)
M (Trust) N (Admire)
Analytically Trained MDs
H (Reject)
J (Attack) P (Resentful) I (Punish)
R (Distrust) Q (Complain)
D (Support) E (Love)
C (Give) M (Trust) L (Cooperate)
N (Admire) K (Appreciates)
MDs without Analytic Training
I (Punish) J (Attack)
H (Reject) P (Resentful) F (Boss)
R (Distrust) Q (Complain)
C (Give) N (Admire)
D (Support) E (Love) K (Appreciates)
M (Trust) O (Conform)
Social Workers
F (Boss) Q (Complain)
I (Punish) J (Attack) H (Reject)
R (Distrust) G (Rebel)
Description of Analytic Patient's Verbalizations
Description of Therapy Patient's Verbalizations

FIG. 7. Stability of meaning in empirical definitions of the positive-negative dimension.

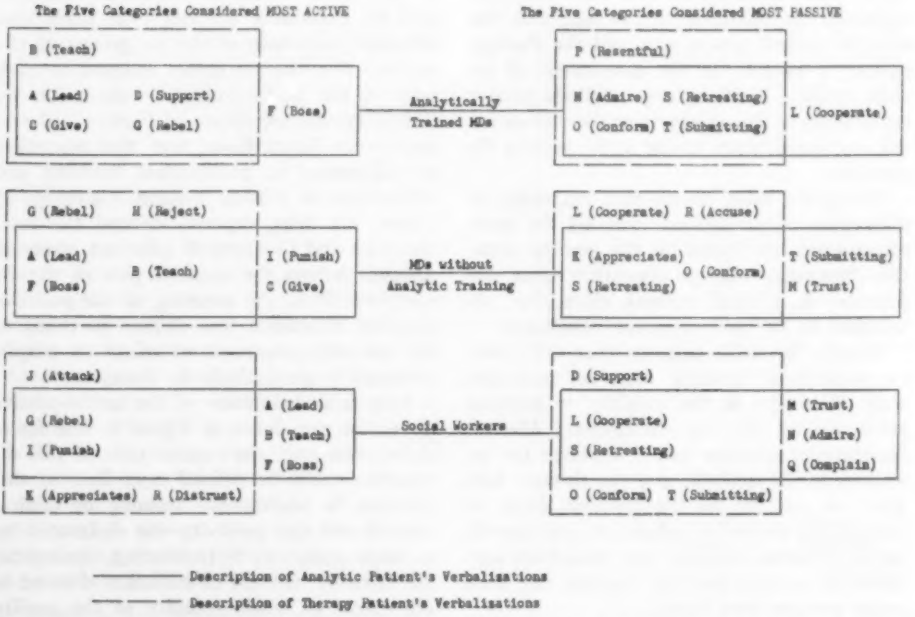
The Five Categories Considered MOST ACTIVE
The Five Categories Considered MOST PASSIVE
B (Teach)
A (Lead) D (Support) F (Boss)
C (Give) G (Rebel)
Analytically Trained MDs
P (Resentful)
N (Admire) S (Retreating) L (Cooperate)
O (Conform) T (Submitting)
G (Rebel) H (Reject)
A (Lead) I (Punish)
F (Boss) B (Teach) C (Give)
MDs without Analytic Training
L (Cooperate) R (Accuse)
K (Appreciates) O (Conform) T (Submitting)
S (Retreating) M (Trust)
J (Attack)
G (Rebel) A (Lead)
I (Punish) B (Teach)
F (Boss)
K (Appreciates) R (Distrust)
Social Workers
D (Support)
L (Cooperate) M (Trust)
S (Retreating) N (Admire)
Q (Complain)
O (Conform) T (Submitting)
Description of Analytic Patient's Verbalizations
Description of Therapy Patient's Verbalizations

FIG. 8. Stability of meaning in empirical definitions of the active-passive dimension.

affected by shifts in clinical context. Of the 10 categories which, in the responses of each group, were assigned the highest active or passive factor loadings, 8 identical categories were used by the criterion group of analytically trained psychiatrists in the description of each case, 6 by psychiatrists without training, and 4 by social workers. The relative immunity of the definition obtained from analytically trained psychiatrists to shifts in clinical context involving sex differences suggests that training in analytic theory (particularly training in the theory of psychosexual development) may contribute to a more generally applicable conceptualization of the active-passive dimension.

In reviewing these findings it should be recalled that whereas the first hypothesis concerning invariant characteristics of clinical thinking postulates that clinicians think alike in that they share a common frame of reference, the second hypothesis postulates that modifications in this common frame of reference may be associated with differences in professional training or differences in the patient described. The first hypothesis was supported by our findings wherever there was high correlation between factor loadings for the two-factor systems produced by each of the groups. When such high correspondence is found in the meaning of the two dimensions, we can have considerable faith in quantitative differences between the groups in terms of the over-all importance assigned to the two-factor system and the relative importance of the two factors. In other words, the second hypothesis concerning modifications of the two-factor system associated with professional training or patient described can be explored in quantitative terms, and qualitative differences should not receive too much attention. Where the first hypothesis concerning the invariant structure of clinical thinking found only moderate support in that correlations between factor loadings for the two factors were relatively low, quantitative comparisons of groups in terms of the over-all importance of the two-factor system or the relative importance of the two factors in each group rest on insecure ground. In this case qualitative modifications of the system, e.g., in the social workers' definition of activity for the female patient, should receive greater attention.

SUMMARY

The purpose of this investigation was to examine the influence of professional training and clinical material upon the meaning of concepts used in case description. Three hundred and one Freudian clinicians were asked to describe the patients' verbalizations in the transcripts of one psychoanalytic and one psychotherapy session by means of 20 randomly presented categories of interpersonal experience. These categories had, in earlier research, been classified in terms of two dimensions derived from analytic theory: positive-negative and active-passive. Responses obtained from these subjects were treated by six separate factor analyses, one each for responses to the analytic and the therapy transcripts by psychiatrists with analytic training, by psychiatrists without analytic training, and by social workers.

Results included the following findings:

1. The two-dimensional system developed a priori was found to characterize clinical case description in that the first two factors underlying the responses of each group represented the postulated dimensions.

2. The two-dimensional system appeared to be more descriptive of the psychotherapy patient than of the analytic patient.

3. In the description of the psychotherapy patient the positive-negative dimension received more and the active-passive dimension received less emphasis than in the description of the analytic patient.

4. Clinicians differed according to their professional training in the relative importance attributed to the two dimensions. In the context of the analytic data social workers responded more as psychoanalysts, while in the context of the therapy data they responded more as psychiatrists without analytic training.

5. The clinicians differed in the extent to which the meaning attributed to the two dimensions remained stable regardless of the patient described. The responses of psychoanalysts showed the most stability of the three groups.

6. The over-all meaning of the positive-negative dimension was found to be more independent of shifts in clinical context or of differences in professional training of subjects than that of the active-passive dimension.

7. Whereas the positive-negative dimension, regardless of differences in professional training or clinical context, was defined in terms of the polarity trust-distrust, definitions of the active-passive dimension varied with professional training and clinical context, particularly with respect to the active pole of the dimension.

REFERENCES

CHANCE, ERIKA. *Families in treatment*. New York: Basic Books, 1959.

CHANCE, ERIKA, & ARNOLD, J. The effect of professional training, experience, preference for a theoretical system upon clinical case description. *Hum. Relat.*, 1960, **13**, 195–213.

FREUD, S. *The standard edition of the complete psychological works of Sigmund Freud*. Vol. XIV. (Originally published 1914–1916) London: Hogarth, 1957. P. 127.

LEVY, L. H., & DUGAN, R. D. A constant error approach to the study of dimensions of social perception. *J. abnorm. soc. Psychol.*, 1960, **61**, 21–24.

OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: Univer. Illinois Press, 1957.

SCHAEFER, E. Converging conceptual models for maternal behavior and for child behavior. In J. Glidewell (Ed.), *Parental attitudes and child behavior*. Springfield, Ill.: Charles C Thomas, 1961, in press.

TUCKER, L. Factor analysis of double centered score matrices. Princeton: Educational Testing Service, 1956.

(Received April 3, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 407–417

EGO INVOLVEMENT:

A CRITICAL EXAMINATION OF SOME METHODOLOGICAL ISSUES

EVA DREIKURS FERGUSON[1]

Craig House for Children, Pittsburgh, Pennsylvania

As psychologists have increasingly turned to the laboratory investigation of human motivation, a host of empirical and conceptual problems has arisen. At the present time, various operational methods for obtaining different degrees of motivation exist, and a variety of types of motivations are being investigated, e.g., anxiety, need achievement, failure-stress, ego involvement. One method of obtaining high and low degrees of a given motivation is to use subjects who differ in their degree of this motivation prior to entry into the laboratory. Another method is to induce different degrees of motivation in the laboratory, in subjects whose existent level of the motivation under study was not ascertained. Although a combination of these methods may be utilized, the use of the latter method has been commonly employed to elicit ego involvement. By appropriate instructions or other means of orienting subjects to the experimental situation, different degrees of an ego involvement motivation are assumed to be established. However, unless a direct measure is made of the extent to which instructions in fact evoke ego involvement motivation, data relating differences in instructions to differences in behavior on performance or learning tasks cannot be readily interpreted. A lack of performance effects from ego involving instructions may be due to the fact that ego involvement motivation was not altered by the instructions, or it may be due to the fact that although instructions produced a change of motivation, this latter change did not lead to the predicted performance changes. Additional problems become apparent from a review of the literature on ego involvement. Stimuli used to evoke ego involvement (EI_S) are not uniform. Tasks presented to subjects and performance measures used to test the effects of EI_S vary greatly. Although some attention has been paid to exploring the stimulus dimension of EI_S (e.g., Shedd & Angelino, 1952), relatively little work has been done in this area. And as yet there appears to be no direct response measure for ego involvement (EI_R), which would eventually allow assessing the degree to which various EI_S in fact evoke the assumed ego involvement motivation.

At least in part as a result of these factors, data on the effects of ego involvement on behavior are inconclusive. Some investigators have shown ego involvement to affect learning (Alper, 1946; Heyer & O'Kelly, 1949) while others have shown it not to do so (Kausler, 1951; Russell, 1952). Where performance has been affected by ego involvement it is sometimes improved (e.g., Kausler, 1951) and sometimes deteriorated (Castaneda & Palermo, 1955; Romanow, 1958; Shedd & Angelino, 1952). Other studies show ego involvement to have no effect (e.g., Irwin & Mintzer, 1942), or to have only a minimal effect (Alper, 1957; Holt, 1945; Vogel, Raymond, & Lazarus, 1959).

A close examination of certain important issues may lead to an ultimate unraveling of the inconclusive nature of present findings in the ego involvement literature. Whether ego involving instructions alter motivation has been answered in the affirmative by two studies, even though neither study used a direct measure specific to the ego involvement motive. In one study, the operational induction of ego involvement was accomplished through

[1] The initial phase of this investigation was begun while the author was a postdoctoral fellow in the Psychology Department, Western Psychiatric Institute, University of Pittsburgh. Grateful acknowledgments are extended to Arthur R. Thomas and Robert E. Rankin of West Virginia University for their invaluable help in serving as experimenters for some of the studies and in their reading of the manuscript. Appreciation also is extended to Robert Eigenbrode for permission to use the performance items of the Rating Test, which he standardized as part of his work in a graduate psychology seminar.

both a preselection of subjects as well as the administration of appropriate instructions (instructions were high or low ego involving in relation to whether subjects were high in n Achievement or high in n Affiliation). And whereas overt performance on a task was barely affected by the instructions, a more direct measure of motivation, i.e., PGR, did show effects, albeit weak effects, from the ego involving instructions (Vogel et al., 1959). In another study, the use of a fairly direct measure of motivation (level of aspiration) showed that when subjects were forced to choose between two responses, clear-cut effects from ego involving instructions were obtained (Feather, 1959). A quite different conceptual approach that may permit a more thorough exploration of the role of ego involving instructions comes from the field-theoretical writings of Wertheimer (1959). He suggests that it is easier to remember relationships, or relational qualities, than absolute qualities of stimuli. Although he referred to memory and to problems of discrimination learning, this suggestion may be relevant to the present problem. If suitably generalized, the suggestion would be that people respond more easily to relational than absolute qualities; as applied to the present problem, one would predict that maximum effects from ego involving instructions will be obtained when subjects respond to the difference between two instructions, instead of responding to merely a single ego involvement instruction as in the more usual experimental procedure.

In line with various of the considerations raised above, the present investigation seeks to delimit some of the characteristics of the ego involvement variable so as to render it more useful for experimental research. Specifically, the present series of studies attempts to explore some important aspects of the stimulus characteristics of ego involving instructions, to tie these where possible to characteristics of ego involvement motivation, and in turn to relate these to a direct response measure of the ego involvement motivation. In this multifaceted aim, emphasis is placed on exploring the degree of, and conditions for, the effectiveness of ego involving instructions, and to delineate the valence characteristics of the ego involvement motivation as these pertain both to the instructions and to the response measures employed in this investigation. Specifically, the following questions were examined.

1. Although several means are possible to test the effectiveness of ego involving instructions, one test for their effectiveness is to establish that ego involving instructions constitute a discriminable stimulus dimension. That is, the majority of subjects when given both the high and the low ego involving instructions should be able to make a reliable distinction between them.

2. Successful performance on a high ego involvement task is more rewarding, or has a higher valence, than on a low ego involvement task. This is an assumption generally held for the ego involvement motivation and at times has been made the operational basis for defining ego involving tasks (e.g., Bayton, 1943). But this assumption has never, to the author's knowledge, been tested directly. That is, no data appear to exist which show that a given high level of performance is more desirable on the high ego involvement task than on the low.[2] A slight extension of this hypothesis would also predict that subjects strive for a higher level of success on the high as against the low ego involvement task (i.e., set higher levels of aspiration for the high as against the low ego involvement task). Evidence for this has been obtained in a study with children (Feather, 1959).

3. Ego involving instructions are far more effective when subjects are in a stimulus-difference set. That is, when subjects have a set of contrasting one ego involvement level to another level, more clear-cut effects from ego involving instructions are obtained than when subjects perceive the instructions as merely an isolated, single event.

4. If under conditions of high ego involvement subjects have an increased desire to do well on the task, one direct measure of the

[2] Since level of aspiration is considered to be a function of both the valence and the probability of attainment of a given goal (Lewin, Dembo, Festinger, & Sears, 1944; Siegel, 1957), level of aspiration data do not provide a direct test of the valence assumption made in Prediction 2. The Transitivity Test devised for the present investigation is an attempt to arrive at such a direct test.

subjects' degree of ego involvement in the task would be "interest in doing well on a task." To establish the validity of this as a response measure of ego involvement (EI_R), two essential criteria must be met: subjects should state higher "interest in doing well on a task" on the high than the low ego involvement task; and the degree of interest in doing well on an ego involving task should remain constant regardless of how well in fact the subject does perform (except at the extreme limits of performance), since this interest is defined by the nature of the task (i.e., by the potential rewards of success on the task) rather than by actual success. This latter criterion distinguishes "interest in doing well on a task" from level of aspiration as a direct response measure of ego involvement, since level of aspiration has been both hypothesized and empirically found to vary with actual performance level on a task (e.g., Lewin et al., 1944; Steisel & Cohen, 1951).

5. A relevant internal drive induced in conjunction with ego involving instructions increases the effect of the ego involving instructions.

METHOD

To test these predictions, and to clarify the extent to which ego involving instructions are effective in eliciting response differences, one major response measure used in the following series of studies was level of aspiration (LA). The author's own experience with LA (Ferguson, 1958) indicated this to be a sensitive measure of motivation, and recent data (Feather, 1959) reversed earlier findings which had failed to show significant effects of ego involvement on level of aspiration. The wording used for LA in this series of studies was such that subjects stated *desired* rather than *anticipated* success, i.e., subjects stated level of aspiration (LA) and not level of expectation (LE). This was necessary to test the hypothesis concerning success preference for the High as against the Low Ego Involvement condition. In line with this, the response measure of ego involvement (EI_R) explored in this investigation was "interest in doing well on the task" instead of "interest in doing the task." For the latter is presumably more contingent on actual than desired success, and also has no necessary relationship to an activity involving the subject's status.

The ego involvement instructions used in this investigation were selected to represent the more typical wording of the two extreme levels of ego involvement instructions most commonly found in the experimental literature. The purported difference between High and Low Ego Involvement instructions is that the subject's status (self-esteem and/or social esteem) is at stake in the former type of task but not, or only minimally, in the latter type of task.

Procedure

Tasks. Two main tasks were used in the present series of studies. On the one task, the Rating Test, a subject received only a single ego involvement instruction; on the other, the Transitivity Test, the subject received both ego involvement instructions. The studies varied with regard to whether one, or the other, or both tasks were given to the subjects. The rationale for administering the experimental tasks in any given study was dependent on the specific hypothesis being tested by that study. As will be seen, the findings reported in the subsequent section are clearly related to the specific task procedure employed.

Rating Test. This consisted of a small five-page booklet. On the first, third, and fifth pages, subjects had merely to make a rating from 0 to 50. On the second and fourth pages, subjects made no ratings but filled in performance items. The Rating Test thus consisted of three rating trials and two performance trials. Each performance trial contained 50 digit substitution items, patterned after the Wechsler Adult Intelligence Scale. Successful and Failing performance was predetermined by alloting subjects enough time to complete approximately 70% of the items per trial in the Success condition and approximately 40% of the items in the Failure condition.

In addition to a rating scale, the first page contained ego involvement instructions. A given subject received only the High *or* the Low instructions. High Ego Involvement instructions stated:

> You are being asked to take a brief intelligence test. Results on this test give a very good indication of a person's intelligence and of the likelihood of career success.

The instructions for Low Ego Involvement were:

> You are being asked to take a test which is in the process of being developed as part of a large experiment. Since you are the first group to do this task, individual scores cannot be adequately evaluated, but your performance will provide some valuable information on what the next steps should be in the further development of this task.

The Rating Test was used to obtain one of two responses. Level of aspiration (LA) was one, the other was an ego involvement response (EI_R) designed especially for this investigation. For LA, the subject placed a check on a scale from 0 to 50 to indicate the number of items he would like to complete in the alloted time. For EI_R, the subject placed a check on a scale from 0 to 50 to indicate how interested he was in doing well on the task. Each subject made only LA *or* EI_R ratings. The ratings on the first page were made before performance was begun, only the ego involving instructions being provided as basis for the subjects' ratings. Ratings on

the third page, following the first performance trial, referred to the subsequent performance trial: how many items the subject would like to complete on the remainder of the task (LA) *or* how interested the subject was in doing well on the remainder of the task (EI_R). Ratings on the fifth page (LA *or* EI_R) referred to another task like the one the subject had just completed but which he was not going to be asked to do at the present time.

Transitivity Test. This consisted of one sheet that contained both ego involvement instructions already described, plus 15 pairs of items. Unlike the Rating Test which provided the subject with only one ego involvement instruction, the Transitivity Test forced the subject to contrast one instruction with another. This test was designed not only to ascertain the subjects' responses in a stimulus-difference situation, but also to test the transitivity and direction of the subjects' success preferences for the two ego involvement conditions. (Transitivity exists when A is preferred to B and B is preferred to C, and A in turn is preferred to C.)

On the top of the sheet were two paragraphs, each paragraph giving the same ego involvement instructions provided on the Rating Test. For control, half the sheets had the High Ego Involvement instructions for the first paragraph, the other half had the Low Ego Involvement instructions for the first paragraph. The High Ego Involvement task was labeled Intelligence Test and the Low Ego Involvement task was labeled Experimental Test. Subjects then had to make 15 pair comparisons, made up of six units: three goals (50, 40, and 30 items completed on a task consisting of 50 items), for the two ego involvement tasks. Each unit was compared once to every other unit, and subjects had to check the preferred item for each of the 15 pairs of items. An illustrative example was provided:

(*a*) 30 on Experimental Test ———
(*b*) 30 on Intelligence Test ———

If you have checked (*a*), that means you would prefer to complete 30 items on the Experimental Test in contrast to 30 on the Intelligence Test.

Subjects

Subjects were drawn from three institutions: University of Pittsburgh (introductory sociology and introductory psychology), West Virginia University (introductory psychology), and Pennsylvania State University (first year student nurses who were in a 3-year hospital program). Subjects took the experiments in their regular class time, without prior knowledge that they were going to be subjects, i.e., they were a captive audience and not volunteers. The series of studies covered a 2-year period, but no subject served in more than one study. Assignment of subjects to the experimental conditions within each study was such that any effects of the experimental variables were not due to systematic subject bias. Tests of significance were made only between subjects serving in each self-contained study. Where the N within cells was not equal, a nonparametric analysis of variance which yields chi square (Wilson, 1956) was used; where N within cells was equal, an F test was used to test the significance of the data.

RESULTS

The data indicate several clear-cut findings:

1. Ego involvement instructions are effective but weak; i.e., they yield clear-cut effects, but only under certain conditions. And though they are reliably distinguished by a majority of subjects, they fail to be so distinguished by a sizable minority of subjects.

2. The valence of success is greater for High than Low Ego Involvement tasks.

3. Instructions yield clear-cut and consistent effects only when subjects define the ego involvement conditions in a context of contrast.

4. Even with the addition of a relevant internal drive, ego involving instructions fail to produce clear-cut and consistent effects when subjects are not in a stimulus-difference set (i.e., when not defining the ego involvement conditions in a context of contrast).

5. Interest in doing well on a task (EI_R) is a valid and reliable measure of ego involvement, but it is sensitive to ego involving instructions only for selected subjects.

Evidence for the effectiveness of ego involvement instructions comes from two main sources: on the Transitivity Test the majority of subjects were able to make consistent goal preference discriminations between tasks differing in degree of ego involvement; when subjects were in the appropriate set, their responses on the Rating Test showed significant differences between High and Low Ego Involvement conditions. Four groups, or a total of 197 subjects, took the Transitivity Test (Table 1). In each group, the majority of subjects, ranging from 54% to 69%, made a reliable goal preference distinction between the High and Low Ego Involvement tasks. The nature of this distinction was a highly significant preference for success on the High as against the Low Ego Involvement task. The vast majority of subjects who made transitive choices did so in the predicted direction, i.e., they chose 50 on an intelligence test as their most preferred goal, with 40 on

TABLE 1

CONSISTENCY OF PREFERENCES FOR PERFORMANCE GOALS ON AN INTELLIGENCE TEST AS COMPARED TO AN EXPERIMENTAL TASK

	Student nurses (N = 33)	University of Pittsburgh[a] (N = 49)	University of Pittsburgh[a] (N = 53[a])	West Virginia University (N = 62)
Number of subjects making transitive choices	18	34	34	41
Number of subjects making transitive choices in the predicted direction[b]	14	24	26	31

[a] Although the total N in this group was 54, the Transitivity Test was not handed in by one subject in that group.

[b] With six goal units in any possible combination of ranks, the chance probability of getting the combination of "50 on the Intelligence Test" in first rank and "40 on the Intelligence Test" in either second or third rank is 1 in 15 (48/720). Using the standard error of a proportion (where the obtained proportion is, e.g., 14 out of 18, and the chance proportion is 1 out of 15), the number of subjects whose transitive choices are in the predicted direction far exceeds the .01 level of significance for each of the four groups.

an intelligence test as their second or third preference.[3] Clearly, when ego involvement instructions constitute a discriminable dimension for subjects, their preference is very strongly in favor of success on the High Ego Involvement task. This same preference direction was seen in the subjects' responses on the Rating Test. When clear-cut ego involvement effects were obtained, the LA and EI_R ratings were significantly higher (higher goals desired, and greater interest in doing well) for the High than the Low Ego Involvement conditions.

Of interest is the fact, however, that ego involvement effects *failed* to occur unless subjects were in the appropriate set. Only when subjects were aware of a contrast between two ego involvement stimuli did their responses show a clear-cut effect from the instructions. It will be recalled that on the Transitivity Test subjects were responding concurrently to two ego involvement instructions, but on the Rating Test they were exposed to only one specific instruction. When the Rating Test *followed* the Transitivity Test, subjects' ratings showed significant differences between the two ego involvement conditions (Table 2). When the Rating Test *preceded* the Transitivity Test, no significant effects from the ego involvement conditions were obtained, and this was true even for those subjects who had a clear-cut preference in favor of success on the High Ego Involvement task (Table 3). Thus in spite of subjects' inherent preference for success on the one as against the other task, without an appropriate set subjects' goal setting behavior fails to reflect this preference. This points to the conclusion that without an appropriate set, subjects' responses were not based primarily on the ego involving instructions but on some other stimulus variables outside of the experimenter's control and manipulation.

The evidence thus far shows that ego involvement instructions do form a discriminable dimension, at least for a majority of subjects, and this dimension has a valence scale such that success is preferred on the High Ego Involvement task. But ego involvement instructions are not strong stimuli. This is seen not only by the large number of subjects who failed to make transitive choices between the two ego involvement tasks, but also by the fact that when subjects were given only a single ego involvement condition without an appropriate set, consistent and clear-cut ego involvement effects failed to appear. This was found so unvaryingly that it cannot be dismissed as merely "negative effects." Two studies were conducted without the use of the Transitivity Test. In the first study, subjects took the experiment early in the semester and the experimenter (a woman) was a stranger to the subjects; in the second study the experiment was conducted one week before final exams and the

[3] Subject's choices were said to be transitive when the six goal items had an internally consistent rank order. That is, the item in first rank was preferred to all other items with which it was paired; the second rank item was preferred to all items with which it was paired *except* for the item in first rank (for that pair of items, the first rank item was preferred over the second rank item). This consistency of choices had to hold for all six goal items, for the 15 pair comparisons, in order for a subject's choices to be called "transitive." The two most common sequences of transitive preferences were the following: 50 Intelligence, 40 Intelligence, 30 Intelligence, 50 Experimental, 40 Experimental, 30 Experimental; 50 Intelligence, 50 Experimental, 40 Intelligence, 40 Experimental, 30 Intelligence, 30 Experimental.

TABLE 2

EFFECTS OF EGO INVOLVEMENT (EI) ON LEVEL OF ASPIRATION (LA) AS A FUNCTION OF SUBJECTS' SET

Trial	When LA precedes the Transitivity Test (University of Pittsburgh$_A$)					When LA follows the Transitivity Test (University of Pittsburgh$_B$)				
	High EI (N cell = 22)		Low EI (N cell = 27)		χ^2	High EI (N cell = 27)		Low EI (N cell = 27)		χ^2
	M	SD	M	SD		M	SD	M	SD	
First (Preperformance)	43.9	9.3	42.9	9.8		45.3	8.4	38.1	8.2	
Second (After Failure)	39.6	9.3	36.6	12.5		37.1	12.9	28.3	11.5	
Third (After Failure)	39.6	8.9	30.8	16.9		36.4	13.3	26.1	12.7	
Between High-Low EI					.14[a]					16.71**

[a] Although a significant difference between the two levels of ego involvement exists for the third trial ($t = 2.33$, $p < .05$), the combined results for all three trials show no significant effect from the Ego Involvement variable.
** $p < .01$. The chi square is based on Wilson's (1956) nonparametric analysis of variance, with 1 df for a High vs. Low EI comparison.

experimenter (a man) was the subjects' own instructor. Both studies showed that when subjects were not in the appropriate set, effects of ego involvement instructions failed to be significant even though subjects responded to Success and Failure as predicted (Tables 4 and 5). That is, LA was significantly lower under conditions of Failure, and EI_R, as hypothesized, was stable with performance. The lack of clear-cut LA and EI_R differences for ego involvement conditions must be assumed to be due to the ineffectiveness of the ego involvement instructions, since LA was extremely sensitive to performance and EI_R was highly reliable. (For the groups cited in Table 4, four correlations for EI_R were computed, for first vs. second trial and first vs. third trial on the combined Success and Failure subjects. These correlations ranged from .77 [$N = 63$] to .89.)

It could be argued that the lack of effects from instructions, shown in Tables 4 and 5, is due to the fact that the High Ego Involvement instructions were not ego involving enough and that the Low Ego Involvement instructions were too ego involving (i.e., not low enough in ego involvement). To test this, two experiments were conducted so that ego involvement would be both raised and lowered. An initial study explored not only the role of a relevant drive (subjects taking the experiment immediately prior to a course ex-

TABLE 3

EGO INVOLVEMENT EFFECTS AS A FUNCTION OF SET AND PREFERENCE[a]

Rating	When ratings preceded the Transitivity Test					When ratings followed the Transitivity Test				
	LA (University of Pittsburgh$_A$)					LA (University of Pittsburgh$_B$)				
	High EI (N = 5)		Low EI (N = 12)		χ^2	High EI (N = 8)		Low EI (N = 11)		χ^2
	M	SD	M	SD		M	SD	M	SD	
First rating	46.2	8.5	48.2	3.4		49.8	.7	40.9	7.7	
Second rating	39.0	10.2	40.2	13.3		41.9	8.4	25.9	12.0	
Third rating	38.5	10.6	35.1	17.2		38.1	11.9	25.8	13.6	
Between High-Low EI					.26					12.45**

[a] When subjects' transitive choices clearly favored success on the High EI task. Because only a very limited number of subjects strongly favored success on a Low Ego Involvement task, the present comparisons are based only on ratings of subjects who strongly favored success on a High Ego Involvement task, i.e., subjects whose transitive choices were "50 and 40 on an Intelligence Test" as two of the top three ranks and "50 on an Experimental Test" as no higher than third rank.
** $p < .01$.

TABLE 4

EFFECTS OF EGO INVOLVEMENT AND SUCCESS–FAILURE ON EI_R AND LA

Trial	EI_R						x^2	LA						x^2
	High EI			Low EI				High EI			Low EI			
	N	*M*	*SD*	*N*	*M*	*SD*		*N*	*M*	*SD*	*N*	*M*	*SD*	
First														
Success	22	36.9	12.1	32	38.5	10.3		36	40.2	9.2	38	36.4	12.3	
Failure	32	41.8	6.7	31	40.3	11.4		32	37.5	10.1	29	36.1	9.5	
Second														
Success		38.3	12.7		39.4	11.0			42.3	8.4		39.1	10.2	
Failure		42.4	10.1		43.2	11.0			29.8	9.5		26.8	10.3	
Third														
Success		36.1	15.3		38.6	11.5			41.8	8.4		39.4	9.1	
Failure		43.6	7.5		41.7	11.8			25.5	11.8		22.5	11.9	
First														
Between High-Low							.01							.60
Between Success-Failure							1.85							.14
Second														
Between High-Low							.30							2.13
Between Success-Failure							1.67							35.00**
Third														
Between High-Low							.24							.07
Between Success-Failure							3.88[a]							55.98**

[a] Although $p < .05$, the effect of performance on EI_R fails to be significant when the Success-Failure differences on the first (preperformance) trial are taken into account. When the change scores (from the first to the third trial) of the two combined Success groups were compared with those of the combined Failure groups, the obtained t of 1.35 fell short of the 2.00 value required for a .05 level of significance (with 62 and 53 *df* and heterogeneous variances). Thus the hypothesis is supported that differential success does not have a significant effect on EI_R.

** $p < .01$ (chi square at the .01 level is 6.64).

amination vs. taking the experiment in a relaxed and informal atmosphere) but also whether anonymity under low ego involvement would help produce the expected ego involvement effects. If these two added variables would be effective in eliciting significant differences from the ego involving instructions, the predicted differences would be

TABLE 5

EFFECTS OF EGO INVOLVEMENT AND FAILURE WHEN SUBJECTS[a] ARE IN AN EGO INVOLVED RELATIONSHIP WITH THE EXPERIMENTER
(*N* per cell = 26)

Trial	EI_R				*F*	LA				*F*
	High EI		Low EI			High EI		Low EI		
	M	*SD*	*M*	*SD*		*M*	*SD*	*M*	*SD*	
First (Preperformance)	39.8	6.0	39.1	7.9		42.0	9.5	43.4	6.3	
Second (Failure)	38.8	8.5	41.0	8.0		36.0	9.7	32.4	11.4	
Third (Failure)	37.5	11.2	39.4	10.4		33.4	13.3	28.4	13.7	
First vs. Second										
Between High-Low EI					.14					.25
Between Neutral-Failure					.27					41.18**
First vs. Third										
Between High-Low EI					.07					.49
Between Neutral-Failure					.58					42.33**

[a] Subjects are from West Virginia University.

** $p < .01$. Using the analysis of variance design suggested by Edwards (1950) for repeated measurements of the same subjects, the *F* ratios represent 1, 50 *df*. (*F* at the .01 level is 7.17.)

TABLE 6

EFFECTS OF EGO INVOLVEMENT ON LA WHEN A RELEVANT DRIVE AND ANONYMITY UNDER LOW EI ARE ADDITIONAL VARIABLES

Trial	First experiment (Student nurses)				F	Second experiment (West Virginia University)				χ^2
	High Drive (N cell = 12)		Low Drive (N cell = 12)			High Drive (N cell = 21)		Low Drive (N cell = 20)		
	M	SD	M	SD		M	SD	M	SD	
First (Preperformance)										
High EI	43.1	7.2	39.6	9.3		40.8	9.5	46.6	5.4	
Low EI	40.9	8.8	36.5	7.4						
Low EI-Anonymity	41.9	6.5	38.2	8.2	.58	46.2	7.9	38.6	12.0	
Between EI conditions										.78
Second (Failure)										
High EI	35.2	11.1	39.6	9.0		34.7	8.4	42.2	10.6	
Low EI	32.0	10.3	32.1	8.9						
Low EI-Anonymity	28.2	7.3	33.5	5.2		37.9	11.8	29.9	12.4	
Between EI conditions					3.34					3.12
Third (Failure)										
High EI	38.1	11.1	43.9	6.7		34.9	11.7	42.9	11.3	
Low EI	30.0	11.2	30.1	10.7						
Low EI-Anonymity	25.8	4.2	32.5	9.8		37.1	15.2	28.8	12.1	
Between EI conditions					10.98**					3.96*

* $p < .05$. Though instruction effects occur only on the third trial, Instruction × Drive interaction is significant for all three trials, in opposite direction to that predicted.
** $p < .01$.

in the direction of highest LA for the High Drive-High Ego Involvement condition and lowest LA for the Low Drive-Low Ego Involvement-Anonymity condition, with the remaining four conditions yielding appropriate LAs between these two extremes. The predicted effects were not obtained, in either the initial or the follow-up study. The latter used only four conditions (Low Ego Involvement being given only under anonymity) to maximize the possibility of obtaining the predicted ego involvement effects within the limit of the available number of subjects. In both studies, ego involvement effects appeared to a significant extent only on the third trial, and then not as predicted (Table 6). Drive effects were not significant in either study, and analysis of variance between the

TABLE 7

EFFECT OF EGO INVOLVEMENT ON EI_R WHEN SUBJECTS[a] ARE IN AN APPROPRIATE SET

Trial	For total group				F[c]	For selected subjects[b]				χ^2
	High EI (N cell = 31)		Low EI (N cell = 31)			High EI (N cell = 12)		Low EI (N cell = 7)		
	M	SD	M	SD		M	SD	M	SD	
First (Preperformance)	44.1	6.4	40.1	10.4		44.8	5.7	36.4	11.3	
Second (Failure)	44.2	6.8	41.6	11.8		45.5	4.6	36.6	13.7	
Third (Failure)	44.1	7.8	41.7	10.0		44.6	6.9	34.7	13.9	
Between High-Low EI					2.36					8.52**
Between trials					.37					.14

[a] Subjects are from West Virginia University.
[b] These subjects strongly preferred success on a High EI task (see footnote to Table 3).
[c] The F tests follows Edwards' (1950) suggested analysis for repeated measurements on the same subjects.
** $p < .01$.

two Low EI conditions in the initial study showed no significant effects for Anonymity.[4] In the second experiment, Drive and Instruction interaction was significant (it was not in the first study), but in the opposite direction from that predicted. These data serve only to confirm the results from the preceding series of studies, and indicate that when subjects are not in an appropriate set, clear-cut and consistent effects from ego involvement instructions fail to occur. Moreover, it does not appear tenable to account for the lack of ego involvement effects by an insistence that the fault lies in the particular instructions used in this series of studies. For not only do these instructions represent the type often used in the experimental literature, but even when the High instructions were augmented by an ego involvement drive and the Low instructions were made minimally ego involving through subjects' performance and ratings being anonymous, the instructions failed to be reliably effective.

One remaining study was carried out. Since a method had been found for eliciting instruction effects, by giving subjects the Transitivity Test first and thereby inducing an appropriate set, a means was now available to test whether EI_R is a suitable response measure for ego involvement. As seen in Table 7, subjects whose preference was clearly in favor of success on the High Ego Involvement task showed a significant difference between the High and Low Ego Involvement conditions. But when the ratings of all subjects are compared, significant ego involvement effects were not obtained even though the Transitivity Test preceded the Rating Test. Thus EI_R is reliable, and a valid measure of ego involvement for selected subjects; but for group data it is not as sensitive to ego involving instructions as is LA. The use of EI_R instead of LA would be preferable, however, in cases where stability of response is desired, since LA does vary with performance and EI_R, as hypothesized, does not do so even when it is sensitive to the ego involving conditions (see chi square for Between trials in Table 7).

[4] Since the conditions of High vs. Low Drive and Anonymity vs. Nonanonymity were introduced into the experiments to heighten and lower ego involvement motivation, they can be viewed as additional categories of EI_s. The negative results obtained from these two studies therefore suggest that the conclusions derived from the studies using only instructions as EI_s may have wider generality.

DISCUSSION

The data that have been presented show clearly that not only are ego involvement instructions less potent than has been generally assumed, but also that they will fail to produce clear and consistent response differences unless subjects define the stimuli as part of a discriminable dimension. Evidence for the weakness of ego involving instructions comes from several independent sets of data, but major evidence for it is the fact that a large minority of subjects do not discriminate reliably between two levels of ego involvement, in terms of the implicit valence scale generally assumed to exist for this variable. An argument against the findings of the present investigation might be raised on the basis that other published studies show clearly significant effects from ego involving instructions without the use of an appropriate prior set. However, the review of the literature already presented should leave little doubt that the evidence regarding the effectiveness of ego involvement instructions is far from consistent. In view of publication policy regarding negative findings, one can only surmise that other investigations which failed to obtain significant ego involvement effects are not lacking. Even among published studies, negative results and data in the direction opposite to that predicted are reported. Since ego involvement has been shown in this investigation to be a reliable phenomenon, even though weak, one can expect that occasional significant results will be obtained without prior set induction.[5]

The findings of the present investigation not only shed new light on the ego involvement variable, but have implications for other motivational variables as well. The data from the present series of studies appear to make the most sense if one assumes that the human

[5] Especially the forced-choice technique used by Feather (1959) would be a likely means of obtaining effects from ego involving instructions, since this at least narrows the response alternatives even if it does not narrow the stimulus field.

adult in most situations is exposed to a great number of internal and external stimuli at any given moment. When, e.g., a single ego involvement condition is presented, subjects are probably exposed to many additional motivations: the need to achieve, the need to attend to the task, the desire to please, the avoidance of failure, and so on. Other stimuli, aside from motivational ones, are also present. (A recent study—Underwood & Postman, 1960—has explored this question in another context.) If the experimental stimulus is already a weak one, as has been demonstrated for ego involving instructions in the present series of studies, the probability that subjects will respond to the experimental stimulus and not to other ones is quite small. However, when an appropriate set is induced, there is both a narrowing of the stimulus field and a specifying of the unique properties of the experimental stimulus. Once the subject is in this set, the probability of his responding to the *unique* stimulus that the experimenter expects him to respond to is greatly increased. In the present investigation, the Transitivity Test was a simple way of inducing such a stimulus set. Without it, subjects failed to respond to the experimental stimulus in terms of its unique properties even though they were able to ascribe these properties to the stimulus. The most striking evidence for this came from the data of the subjects whose choices were transitive and in favor of success on a High Ego Involvement task. Without an appropriate set, their level of aspiration was uniformly high, regardless of the task instructions. This suggests that they were responding to other variables, rather than primarily to the experimental one. But when subjects were in an appropriate set, these same instructions produced responses far more in line with the main experimental condition, the most notable effect being a lowered LA for the Low Ego Involvement condition. This finding of the importance of a relevant stimulus set suggests that possibly other motivational variables could be brought into sharper focus by an appropriate set induction. Studies in other areas of human motivation (e.g., stress, anxiety, n Achievement) have suffered some of the same inconclusive results found in the literature of ego involvement. When motivation in human subjects is experimentally induced, this may indeed elicit only weak stimuli, as critics of this approach to human motivation have been free to point out. Granting the possible limits of laboratory defined motivations, the induction of a relevant set comparable to that employed in the present investigation may make for more effective study of such motivations.

SUMMARY

The present investigation explored some important aspects of the ego involvement variable so as to render it more useful for experimental research. A series of studies revealed the following:

1. Ego involvement instructions are effective but less potent than has been generally assumed in eliciting the desired motivation.

2. To a very significant degree, subjects prefer success on a High as against a Low Ego Involvement task.

3. Consistent and clear-cut effects from ego involvement instructions occur only when subjects are in an appropriate set, of perceiving a given instruction as part of a dimension consisting of contrasting levels of ego involvement. Without such a set, ego involvement instructions are not reliably effective, and this is true even when instructions are combined with a relevant drive and when subjects under low ego involvement take the task anonymously.

4. Interest in doing well on a task is a valid response measure of ego involvement for selected subjects. As hypothesized, it is stable with performance. But for unselected subjects it is not as sensitive to ego involvement instructions as is level of aspiration which is *not* stable with performance.

Limiting the number of cues to which subjects respond in an experiment dealing with motivation appears to be a necessary condition for eliciting significant response differences. The Transitivity Test devised in this investigation contrasted two levels of ego involvement and thus focused the subjects on the specific experimental cues. Through such a technique, which emphasizes relations rather than absolute attributes of experimental instructions, it should be possible to establish relevant orientations for other motivational

variables and thus contribute to added precision in their study.

REFERENCES

ALPER, THELMA G. Task-orientation vs. ego-orientation in learning and retention. *Amer. J. Psychol.*, 1946, **59**, 236–248.

ALPER, THELMA G. Predicting the direction of selective recall: Its relation to ego strength and *N* achievement. *J. abnorm. soc. Psychol.*, 1957, **55**, 149–165.

BAYTON, J. A. Interrelations between levels of aspiration, performance, and estimates of past performance. *J. exp. Psychol.*, 1943, **33**, 1–21.

CASTANEDA, A., & PALERMO, D. S. Psychomotor performance as a function of amount of training and stress. *J. exp. Psychol.*, 1955, **50**, 175–179.

EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart, 1950.

FEATHER, N. T. Subjective probability and decision under uncertainty. *Psychol. Rev.*, 1959, **66**, 150–164.

FERGUSON, EVA D. The effect of sibling competition and alliance on level of aspiration, expectation, and performance. *J. abnorm. soc. Psychol.*, 1958, **56**, 213–222.

HEYER, A. W., & O'KELLY, L. I. Studies in motivation and retention: II. Retention of nonsense syllables learned under different degrees of motivation. *J. Psychol.*, 1949, **27**, 143–152.

HOLT, R. R. Effects of ego-involvement upon levels of aspiration. *Psychiatry*, 1945, **8**, 299–317.

IRWIN, R. W., & MINTZER, MARCIA G. Effect of differences in instructions and motivation upon measures of the level of aspiration. *Amer. J. Psychol.*, 1942, **55**, 400–406.

KAUSLER, D. H. A study of the relationship between ego-involvement and learning. *J. Psychol.*, 1951, **32**, 225–230.

LEWIN, K., DEMBO, TAMARA, FESTINGER, L., & SEARS, PAULINE S. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. I. New York: Ronald, 1944. Pp. 333–378.

ROMANOW, CONCETTA V. Anxiety level and ego involvement as factors in concept formation. *J. exp. Psychol.*, 1958, **56**, 166–173.

RUSSELL, W. A. Retention of verbal material as a function of motivating instructions and experimentally-induced failure. *J. exp. Psychol.*, 1952, **43**, 207–216.

SHEDD, C. L., & ANGELINO, H. R. The effect of ego-involvement on learning. *Proc. Okla. Acad. Sci.*, 1952, **33**, 285–288.

SIEGEL, S. Level of aspiration and decision making. *Psychol. Rev.*, 1957, **64**, 253–262.

STEISEL, I. M., & COHEN, B. D. The effects of two degrees of failure on level of aspiration and performance. *J. abnorm. soc. Psychol.*, 1951, **46**, 79–82.

UNDERWOOD, B. J., & POSTMAN, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.*, 1960, **67**, 73–95.

VOGEL, W., RAYMOND, SUSAN, & LAZARUS, R. S. Intrinsic motivation and psychological stress. *J. abnorm. soc. Psychol.*, 1959, **58**, 225–233.

WERTHEIMER, M. On discrimination experiments: I. Two logical structures. *Psychol. Rev.*, 1959, **66**, 252–266.

WILSON, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, **53**, 96–101.

(Received March 6, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 418–424

PARENTAL ATTITUDES ASSOCIATED WITH SOCIAL DEVIANCE IN PREADOLESCENT BOYS[1]

C. L. WINDER[2] AND LUCY RAU

Stanford University

This is a study of the relationships between parent attitudes and social deviance of preadolescent boys. "Social deviance" is here defined as the extent to which a child's characteristic interpersonal behavior is extreme on one or more of several general dimensions as reported by his peers.

METHOD

Measures

The measure used for social deviance is the Peer Nomination Inventory (PNI). This is a modified "guess who" technique, which is administered to classroom groups. Pupils are asked to nominate members of the class on each of 64 items. Each item refers to a fairly specific aspect of behavior and is phrased in language characteristically used by children of this age in describing each other. The instrument yields a score for each child on five variables: Aggression, Dependency, Withdrawal, Depression, and Likeability. The psychometric properties of the PNI are described by Winder and Wiggins (1960). It should be noted that all the dimensions are intercorrelated (see Table 1). Thus the instrument may be considered a measure of general social deviance which is negatively related to "likeability," or sociometric choice.[3]

We assume that parent characteristics are important antecedents of social deviance in boys of this age. Accordingly we have developed a questionnaire measure of parent attitudes, the Stanford Parent Attitude Questionnaire. The current version of the questionnaire has two forms, one for mothers and one for fathers. They consist of 491 and 518 items, respectively, grouped into 27 and 28 scales. The list of scales, including six summary scales, and the number of items for each are given in Table 2. The majority of the items, except for those in the Parental Adjustment and the Models groups of scales, are identical for mothers and fathers. One scale, Closeness to father, applies to fathers only.[4]

The variables to be measured were selected on the basis of the available studies of parental antecedents of personality development in children. These studies include those of child rearing practices of parents of normal preschool and older children, e.g., Sears, Maccoby, and Levin (1957), and Miller and Swanson (1960). Other researchers, e.g., Bandura (1960), have systematically studied parents of particular groups of deviant children. In addition, our choice of relevant parent characteristics was influenced by clinical experience.

Items were selected on an a priori basis to represent each of the variables. The main sources of items were verbatim transcriptions of interviews with parents of deviant and normal children.[5] These item samples reported actions toward child and others, attitudes toward child, spouse, and self, and beliefs about child rearing and family roles and relationships. Whenever possible the original phrasing of the statements was preserved. The items are thus in the language actually used by parents in reference to particular people and situations, and so differ markedly from the kinds of items used in other parent questionnaires. Items were assigned to scales on the basis of the judgment of the authors and a graduate student in clinical psychology.[6]

The general strategy of the study was to collect responses to the parent questionnaire from a sample of mothers and fathers whose sons had received high, medium, or low scores on the PNI variables, Aggression and Dependency.

Procedure

Sample. The subjects of this study were drawn from a total population of 710 fourth, fifth, and

[1] The study was supported in part by USPHS Grants M-2745 and M-2745 (C1).

The cooperation and assistance of personnel of the Palo Alto Unified School District is gratefully acknowledged. Henry Gunn, M. John Rand, John Caffrey, and H. B. Gelatt were most helpful. Principals and teachers at nine elementary schools gave time and support which was essential.

[2] Now at Michigan State University.

[3] Although the items of the PNI are worded in the language of children of this age, the constructs which the scales purport to measure are in the authors' opinion central to deviant personality development. For the rationale of scale construction see Winder and Wiggins (1960). Validation studies are in progress.

[4] For each scale, the item having the highest correlation with the total scale score has been identified. The list of these items has been deposited with the American Documentation Institute. Order Document No. 7091 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

[5] Interviews from studies by Bandura, by Sears and his associates, and from the Stanford Psychological Clinic were used.

[6] Ronald Cowan.

sixth grade boys in nine Palo Alto schools to whom the PNI had been administered. Cases were selected on the basis of their scores on either the Aggression variable or the Dependency variable. Half the cases were drawn equally from the top, middle, and bottom thirds of the distribution of scores on Aggression; the other half similarly represented the Dependency distribution. Letters were sent to the parents of these children requesting their participation in the study. Of the 262 families contacted, 64% responded, 55% agreed to take the questionnaire, and usable questionnaires were actually obtained from 108 fathers and 118 mothers. Comparisons were made to discover if the sample which took the questionnaire differed either from the total sample or from the sample which was asked to take the questionnaire but did not do so. There are no significant differences on the PNI variables except on Likeability. The critical ratio for the difference between the mean scores on Likeability for the respondents and the

TABLE 1

SCALE RELIABILITIES AND INTERCORRELATIONS FOR THE PEER NOMINATIONS INVENTORY (PNI)

Scale	A	B	C	D	L
A. Aggression	.85				
B. Dependency	.74	.84			
C. Withdrawal	.29	.47	.88		
D. Depression	.48	.62	.62	.77	
L. Likeability	−.23	−.33	−.54	−.41	.94

Note.—N = 710. All correlations significant at $p < .001$.

TABLE 2

STANFORD PARENT ATTITUDE QUESTIONNAIRE VARIABLES, NUMBERS OF ITEMS, AND SCALE RELIABILITIES

Scale	Mothers		Fathers	
	Number of items	Reliability	Number of items	Reliability
I. Ambivalence				
1. Affection demonstrated	37	.66	37	.74
2. Contingent reward	10	.57	10	.50
3. Rejection	39	.66	40	.68
Total	86	.63	87	.67
II. Strictness				
4. Sex anxiety	15	.72	15	.71
5. Aggression anxiety	18	.65	18	.72
6. Restrictiveness	18	.57	18	.56
7. Demands for conformity	21	.62	21	.71
8. Consistency	7	.56	7	.52
9. Conscience	23	.61	23	.62
10. Deprivation of privileges	5	.50	5	.53
Total	107	.86	107	.89
III. Aggression and Punitiveness				
11. Demands for aggression	35	.73	36	.69
12. Punitiveness and physical punishment	11	.42	11	.61
13. Parental aggression	15	.65	15	.47
Total	61	.73	62	.70
IV. Parental Adjustment				
14. Self-esteem	20	.71	24	.68
15. Sociability	11	.52	11	.61
16. Positive father-mother relation	32	.77	46	.73
17. Inconsistency	14	.49	12	.29
Total	77	.67	93	.72
V. Models				
18. Use of models	9	.55	9	.70
19. Sex role expectations	26	.60	23	.45
20. Resemblance to parent	6	.42	10	.49
21. Authority	9	.48	7	.30
22. Closeness to father (fathers only)	0	—	9	.57
Total	50	.48	58	.61
VI. Mastery				
23. Democracy	21	.58	22	.62
24. Nonrestrictiveness	12	.32	12	.37
25. Rewarding independence	25	.69	25	.65
26. Child's mastery	15	.25	15	.40
27. Achievement standards	14	.74	14	.72
28. Reasoning	23	.71	23	.61
Total	110	.87	111	.85

TABLE 3

EDUCATIONAL LEVELS OF PARENTS

Educational level	% Fathers	% Mothers
PhD or MD	7.3	0.0
MA, MS, or graduate work	23.6	7.2
AB or BS	24.5	23.0
Some college	15.5	34.9
High school + technical training	0.9	5.6
High school graduate	19.1	20.6
Less than high school graduate	9.1	5.6

"no shows" is 1.96 ($p = .05$), the boys of the respondents being more popular. The means on the four deviance scales are all slightly higher for the "no shows," but not significantly so. In general we may say that the sample of families where parents took the questionnaire, as judged on the basis of social reputation of the boys, is not severely biased, but does contain a slightly higher proportion of "well-adjusted" boys.

Administration. Questionnaires were administered to groups of parents on several evenings, using the meeting room facilities of the schools. It took most parents about 1.5 hours to complete the questionnaire, an arduous task for a good many parents.

We do not have detailed information on the demographic characteristics of the sample. Palo Alto is a predominantly middle- and upper middle-class suburban community. The educational level distribution of our subjects is shown in Table 3.

Scoring procedure. Each item was answered by checking one of four alternatives: Strongly Agree, Agree, Disagree, Strongly Disagree. A dichotomous method of scoring was used, based on the distribution of responses to each item. A plus score on a given item always indicated presence of the characteristic presumably measured by the scale. Either agreement or disagreement might be scored plus for a given item; a few items were dichotomized for scoring between the Strongly Agree or Strongly Disagree category and all others. Total score for each variable is simply the number of items answered in the scored direction.

TABLE 4

CASES IN HIGH, MIDDLE, AND LOW CATEGORIES ON AGGRESSION AND DEPENDENCY: DISTRIBUTION ON OTHER PNI VARIABLES

Variable A (Aggression)		Variable B (Dependency)			Variable C (Withdrawal)			Variable D (Depression)		
		H	M	L	H	M	L	H	M	L
H	*N*	28	8	0	19	8	9	23	9	4
	%	78	22	0	53	22	25	64	25	11
M	*N*	8	18	9	6	16	13	8	16	11
	%	23	51	26	17	46	37	23	46	31
L	*N*	0	8	25	10	10	13	5	10	18
	%	0	24	76	30	30	40	15	30	55
Variable B (Dependency)		**Variable A (Aggression)**			**Variable C (Withdrawal)**			**Variable D (Depression)**		
		H	M	L	H	M	L	H	M	L
H	*N*	28	8	0	20	7	9	22	9	5
	%	78	22	0	56	19	25	61	25	14
M	*N*	8	17	9	9	14	11	12	16	6
	%	24	50	26	26	41	32	35	47	18
L	*N*	0	8	26	6	14	14	2	10	22
	%	0	24	76	18	41	41	6	29	65

Note.—Ranges of scores for H, M, and L groups were as follows: Variable A: H = 119 and up, M = 33–118, L = 0–32; Variable B: H = 94 and up, M = 33–93, L = 0–32; Variable C: H = 64 and up, M = 13–66, L = 0–12; Variable D: H = 39 and up, M = 14–38, L = 0–13.

TABLE 5

LEVELS OF SIGNIFICANCE FOR DIFFERENCES AMONG MEANS: HIGH, MIDDLE, AND LOW GROUPS

Scale	Mothers					Fathers				
	A	B	C	D	L	A	B	C	D	L
1	.05		.03			.05				
3		.01				.05	.05			
4		.05								
6	.05	.01				.01	.01			
8	.01	.001		.01						
10	.03	.01			.05	.01				
11		.03	.01	.01	.01					.05
12	.03	.01				.001	.01	.03	.01	
14	.05	.05	.05	.05	.03		.05			
19						.03				
26										.01
I	.01	.01	.05	.03		.001	.001		.05	
II		.01				.05	.05			
III		.01	.05	.01	.001	.05	.05			.01
IV					.01					
VI	.05		.05							

Reliability. The reliabilty of each scale was estimated by the Kuder-Richardson procedure (Kuder & Richardson, 1937). These estimates are shown in Table 2. The mean reliability for mother scales and for father scales is .58. No item refinement of these scales has yet been done. The reliabilities are satisfactory for most of the scales, although a few clearly are not. One would expect the reliabilities to increase substantially once further item selection has been done. The mean reliability of the six summary scales is .71 for mothers, .74 for fathers. The total score on Models for mothers is quite unreliable, con-

sidering the number of items; this is also the summary scale with the lowest reliability for fathers.

RESULTS

The first step in the analysis was to divide the sample into a high, middle, and low group for each of the five variables measured by the PNI. Thus each case was assigned to one of three groups on each PNI variable according to the child's score on that variable. The ranges of scores defining the high, middle, and low groups on each variable are shown in Table 4. Table 4 also shows the proportions of the cases assigned to each of the Aggression categories which fall in the respective thirds of the distributions on three other PNI variables. A similar breakdown is given for the Dependency categories. The cutting scores for assigning cases to each group were chosen so as to approximate as nearly as possible a division of the sample into equal thirds. Some consideration also was given to the distribution of scores in the total population from which this subsample was drawn.

On each of the 34 parent attitude scales, a simple analysis of variance was carried out to test the significance of the differences among the mean scores obtained by the parents of the high, middle, and low Aggression cases. This statistical procedure was deemed legitimate since the distributions of scores on the parent scales do not depart markedly from normality and variances of the three groups of cases being compared on each scale are quite similar. Separate analyses were done for mothers and for fathers. The total set of comparisons for all the attitude scales was repeated for each of the other PNI variables. Table 5 presents the results of this analysis. The p values are given for each case in which a parent attitude scale yielded a significant F for the differences in mean scores among the high, middle, and low groups.

From an inspection of Table 5, it appears that 35 comparisons of scores on mothers' scales and 22 comparisons for fathers were significant at the .05 level or better. There were 17 instances in which a parent attitude scale predicted differences in Aggression, 18 for Dependency, 7 for Withdrawal, 7 for Depression, and 8 for Likeability. The cases were originally selected, as described previously, on the basis of their scores on either Aggression or Dependency. No doubt this accounts in part for the greater number of relationships found for these two variables. It will be noted that in 16 out of 26 instances a given attitude scale is associated with more than one PNI variable. This would be expected on the basis of the intercorrelations among the PNI variables.

The relationships found between parent attitude scales and the PNI measures of deviancy (i.e., Aggression, Dependency, Withdrawal, and Depression) may be summarized as follows (the relationships with Likeability are listed separately):

Associated with all four deviancy variables:
- Mothers—Low self-esteem, high ambivalence
- Fathers—High punitiveness and physical punishment

Associated with three deviancy variables:
- Mothers—High consistency, high demands for aggression, high aggression and punitiveness
- Fathers—High ambivalence

Associated with Aggression and Dependency:
- Mothers—High restrictiveness, high deprivation of privileges, high punitiveness and physical punishment
- Fathers—High strictness, high aggression and punitiveness, high rejection, high restrictiveness

Associated with Aggression and Withdrawal:
- Mothers—High affection demonstrated, high mastery

Associated with Aggression alone:
- Fathers—High affection demonstrated, high deprivation of privileges, high sex role expectations

Associated with Dependency alone:
- Mothers—High rejection, high sex anxiety
- Fathers—Low self-esteem

Associated with Likeability:
- Mothers—Low deprivation of privileges, low demands for aggression, low aggression and punitiveness, high self-esteem, high parental adjustment
- Fathers—Low demands for aggression, low aggression and punitiveness, high child mastery

DISCUSSION

The major focus of this study has been on those aspects of parent attitudes found to be associated with general social deviancy in preadolescent boys. If we consider first those attitude scales related to three or all four of

the PNI measures of deviancy, the picture which emerges is one of ambivalence and punitiveness on the part of both parents. In addition, low self-esteem, high consistency, and high demands for aggression are characteristic of the mothers of the generally deviant boys. If we assume that the parent attitudes being measured by this inventory are in fact antecedent to the development of deviant behavior in the children, then we might interpret these findings as indicating that children who experience relatively intense frustration in their interactions with their parents will come to exhibit with considerable intensity a diverse set of maladaptive behaviors. These maladaptive behaviors will ordinarily include aspects of hostile aggression, overdemanding and inappropriate bids for attention, withdrawal from friendly interaction with peers, and such manifestations of sadness and distress as frequent crying.

The measure of "ambivalence," which seems to a highly significant extent to predict deviance in the children, calls for a word of explanation. It is the sum of two scales, rejection and demonstration of affection. Although these two scales correlate slightly negatively in the total population of parents, parents of children who score high on the deviance variables have, on the average, scores above the median on *both* of these scales. What meaning can be assigned to the combined score as a predictor of deviancy? Possibly the parents of the deviant children are acting in both of these somewhat contradictory ways, alternately and inconsistently being affectionate and rejecting toward their sons. It seems more likely, however, that it is primarily the rejection which is related to the deviant behavior in the children, and that these parents' elevated scores on "demonstration of affection" represent a defensive attempt at compensation for their hostile or rejecting attitudes. In such a highly educated, middle-class sample, one would be unlikely to find expressions of rejection without some concomitant manifestations of guilt and defensiveness.

In addition to punitiveness and ambivalence, the other variable most generally associated with deviance is low maternal self-esteem. This finding would seem consonant with the common view that neuroticism, maladjustment, or distress in the mother is likely to prove pathogenic for the child. It seems important, however, to acknowledge the other possible interpretation of these findings, namely, that the deviations in parental attitudes, including lowered self-esteem and rejection of the child, may be the product rather than the antecedent of the child's deviant behavior. Such reactive attitudes of parents may occur where the child's pathology is primarily organic in origin. It is the basic assumption of the present study that parental attitudes may be considered as antecedent to deviant behavior in the children, but obviously neither our data nor those of any other cross-sectional study will provide us with a critical test of this hypothesis.

When we consider the attitude scales associated primarily with Aggression and Dependency, the set of parent characteristics is somewhat amplified but not essentially changed from that associated with deviance in general. It would seem that a combination of restrictiveness, punitiveness, and paternal rejection results in a pattern of hostile and demanding behavior, as reported by peers. There is some agreement with the findings of other studies as to the antecedents of Aggression and Dependency. For example, Sears, Whiting, Nowlis, and Sears (1953) found the highest level of dependency in preschool children to be associated with a combination of reward and punishment, and that aggression increased with punitiveness, up to a certain point. The findings with respect to the effect of permissiveness are contradictory. This may reflect either the difference in ages of the two samples or the more deviant manifestations of Aggression and Dependency being measured in this study.

It is of some interest that a set of father variables appears as specifically related to Aggression and not to the other PNI variables. These scales are father's sex role expectations, demonstration of affection, and use of deprivation of privileges. One might conclude that, given the general conditions for the development of deviance, if the father is interested and involved with his son, actively manipulates rewards and punishments, and is concerned that the son behave in more

masculine ways, the son's deviant behavior is more likely to take a predominantly aggressive form. In contrast, Dependency is especially related to maternal rejection and sex anxiety and low paternal self-esteem. This would suggest that more dependent forms of deviant behavior may develop in a setting in which the father is less confident and assertive and the mother reacts to the child's expression of impulse with anxious concern or aversion.

The parent characteristics associated with high Likeability, which might be considered deviance in the socially desirable direction, are in most instances the reverse of those associated with the other PNI variables. The parents of children who are especially popular and well-liked by their classmates tend not to be aggressive or punitive and have low demands for aggression. The mothers are apparently satisfied with themselves and their relationships with other family members. The fathers express high regard for their sons and evaluate them as competent.

To summarize our conclusions from these data, it would seem that deviant social behavior in boys of this age develops in a setting of punitiveness, restrictiveness, and ambivalence on the part of both parents. Within this general context of disturbed parent-child interactions, we are tempted to speculate on somewhat different roles played by mothers and fathers in the shaping of their sons' deviant behavior. The father seems to be an important influence to the extent that he is actively involved with the child and presents either an assertive and perhaps exaggeratedly masculine or a somewhat unassertive, inadequate model. The mother's impact on the son, on the other hand, seems to be a somewhat more complicated function of her internal dynamics and relationships with other family members, particularly the father. She may react toward the child with hostile rejection or anxious overconcern and restrictiveness at least in part because her own security and self-esteem are impaired. And this, in turn, may be a reflection of a disturbed or unsatisfying marital relationship. Such formulations are only very tentatively suggested by our data, including the intercorrelations among the attitude scales, but they bear some resemblance to Becker's (Becker, Peterson, Hellmer, Shoemaker, & Quay, 1959) conclusions, from his factor analytic study, that the attitudes of mothers and of fathers of child guidance cases were organized somewhat differently.

The sampling procedure followed in this study, as well as the properties of the PNI as a selection instrument, permitted us to study those parent characteristics associated with generally deviant social behavior within the context of the normal male elementary school population. Our findings are thus not directly comparable to studies of child guidance cases or samples of more narrowly selected deviant groups. It is perhaps worth pointing out that the degree of association among the PNI variables, as demonstrated in Table 3, suggests that "pure cases" of a particular form of deviant behavior, such as aggression or withdrawal, are relatively infrequent in the population of deviant children of this age. Perhaps, then, our findings may be considered more representative of the parental characteristics generally associated with socially deviant behavior than those found in studies of more narrowly selected groups. In any case, they are not directly comparable to the findings of studies such as Becker's (Becker et al., 1959) or Bandura's (1960) which specifically contrasted the families of aggressive with those of withdrawn children.

The question might be raised as to the need for another set of parent attitude scales when several others, particularly the PARI (Schaefer & Bell, 1958) and the Shoben (1949), are already quite widely used. Studies by Zuckerman (Zuckerman, Barrett, & Bragiel, 1960) and others (Leton, 1958) have not proved encouraging as to the usefulness of such instruments in attempts to substantiate the connections between parental characteristics and childhood psychopathology. It has seemed to us, moreover, that we might improve on available instruments. Some are too global, and they have not been specifically designed to measure those parent characteristics which have been found in previous studies of normal as well as deviant children to predict differences in specific aspects of personality development. Our choice of variables was obviously determined by a

particular theoretical point of view about personality development as well as by existing empirical evidence. The scales have the additional possible advantage that the items are worded in the natural language of parents. The findings so far would suggest that this approach will prove fruitful.

SUMMARY

Parent attitudes were studied as they related to the social deviance of preadolescent boys. Social deviance is defined by extreme scores on the Peer Nominations Inventory, an instrument which utilizes classmates' ratings. A questionnaire composed of 28 rationally constructed scales referring to aspects of family relationships and child rearing practices was administered to samples of mothers and fathers.

The five social reputation scales of the PNI are Aggression, Dependency, Withdrawal, Depression, and Likeability. The first four are negatively related to Likeability and intercorrelate positively to define a cluster of generally maladaptive social behaviors. Parent attitudes which differentiated deviant from nondeviant boys included ambivalence, punitiveness, demands for aggression, restrictiveness, and low maternal self-esteem. High Likeability was in general associated with the opposite parent characteristics. In addition, the mothers of the popular boys reported high parent adjustment and the fathers gave more favorable evaluations of their sons' competence.

Additional findings suggested that fathers play an important role in the development of deviant behavior in general, but contribute uniquely in several ways to aggression. Fathers of more aggressive boys reported more affection and manipulation of rewards and more stereotyped sex role expectations. A different set of parent characteristics seems specifically related to dependency, namely, mothers' rejection and sex anxiety and fathers' low self-esteem.

REFERENCES

BANDURA, A. Relation of family patterns to child behavior disorders. Progress Report, June 15, 1960, USPHS Grant M-1734.

BECKER, W. C., PETERSON, D. R., HELLMER, L. A., SHOEMAKER, D. J., & QUAY, H. C. Factors in parental behavior and personality as related to problem behavior in children. *J. consult. Psychol.*, 1959, **23**, 107–118.

KUDER, G. F., & RICHARDSON, M. W. The theory of estimation of test reliability. *Psychometrika*, 1937, **2**, 151–160.

LETON, D. A. A study of the validity of parent attitude measurement. *Child Develpm.*, 1958, **29**, 515–520.

MILLER, D. R., & SWANSON, G. E. *Inner conflict and defense*. New York: Holt, 1960.

SCHAEFER, E. S., & BELL, R. Q. Development of a Parental Attitude Research Instrument. *Child Develpm.*, 1958, **29**, 339–362.

SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child-rearing*. Evanston, Ill.: Row, Peterson, 1957.

SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. psychol. Monogr.*, 1953, **47**, 135–234.

SHOBEN, E. J., JR. The assessment of parental attitudes in relation to child adjustment. *Genet. psychol., Monogr.*, 1949, **39**, 101–148.

WINDER, C. L., & WIGGINS, J. S. Measurement of social adjustment in preadolescent boys. Progress Report, November 7, 1960, USPHS Grant M-2745 (C1).

ZUCKERMAN, M., BARRETT, B. H., & BRAGIEL, R. M. The parental attitudes of parents in child guidance cases: I. Comparisons with normals, investigations of socioeconomic and family constellation factors, and relations to parents' reactions to the clinics. *Child Develpm.*, 1960, **31**, 401–417.

(Received April 5, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 425–431

PREFERENCE FOR DELAYED REWARD AS A FUNCTION OF AGE, INTELLIGENCE, AND LENGTH OF DELAY INTERVAL [1]

WALTER MISCHEL
Stanford University

AND

RALPH METZNER
Harvard University

The ability to delay immediate gratification has long occupied a central role in theorizing about child development (e.g., Freud, 1959; Rapaport, 1950; Singer, 1955). Only recently, however, has this type of behavior been subjected to experimental research. One group of studies (e.g., Singer, Wilensky, & McGraven, 1956), using primarily adult schizophrenic subjects, has explored the correlates of "delaying capacity" as reflected in the Rorschach human movement (*M*) response. A second group of studies has approached the problem of delay of gratification directly by eliciting a choice between an immediate, smaller reward (ImR) and a delayed, larger one (DelR). Thus far, using children from other cultures as subjects, preference for DelR has been related positively to social responsibility (Mischel, 1961c), to need for achievement and "naysaying" (Mischel, 1961a), and to the presence of the father in the home under certain conditions (Mischel, 1961b).

The present study attempts to isolate further some of the basic parameters of the preference for delayed as opposed to immediate reward in a simple choice situation. Specifically, three main hypotheses were tested, viz., that preference for DelR should be related (*a*) positively to age and (*b*) to intelligence, and (*c*) negatively to length of the delay interval.

The *first* hypothesis is derived from the psychoanalytic theory of development, according to which the transition from immediate wish fulfilling types of behavior to delayable reality testing types is the crucial element in the change-over from primary to secondary processes. It is assumed that this transition is not an all-or-none affair, but a continuing process resulting in behavior which can be ranged on a continuum from uncontrollable impulsivity to compulsive postponement. The ability to delay gratification in various need areas is presumed to increase with age, and hence in a simple, paradigmatic choice situation older children should more often prefer the delayed, larger as opposed to the immediate, smaller reward. The *second* hypothesis is derived from the same source, more specifically from the psychoanalytic theory of thinking (Rapaport, 1950), according to which learning to delay is intimately bound up with learning to think; i.e., the use of cognitive reality testing replaces the use of uncontrolled motor discharge in the attainment of needs. Hence we would expect greater delaying capacity to be reflected in greater cognitive facility, i.e., intelligence. The *third* hypothesis is a deduction from any expectancy theory which assumes (*a*) that the effective value of an incentive is a joint function of its "real" value and the expectation of getting it, and (*b*) that the expectation of getting something decreases with its temporal distance.

Previous empirical studies, using Trinidadian children (Mischel, 1958) and Palestinian Arab refugee children (Melikian, 1959), have found evidence for a relationship between DelR preference and age, using age ranges from 7 to 9 and from 5 to 10, respectively. The present study extends the age range by using a single measure on children from 5 to 12 years old. Melikian also found a suggestive relationship with intelligence (using Goodenough's Draw-A-Man Test); and Levine, Spivack, and Wight (1959) have summarized a large body of data correlating the Rorschach *M* response, their *indirect* measure of delaying capacity, with intelligence in

[1] This study was supported in part by a PHS research grant (M 3344) from the National Institute of Mental Health, United States Public Health Service, and a grant from the Laboratory of Social Relations, Harvard University. Thanks are also due to Mark Harris for help in data collecting.

adults. It should be noted however that the relationship of *M* to the ability to delay motor discharge has not so far been confirmed for adolescents or children (Spivack, Levine, Fuschillo, & Tavernier, 1959). No previous work has apparently been done on the third hypothesis.

In addition, the present experiment seeks to clarify the relationship of delaying capacity to length of future time perspective. Previous work, using adolescent subjects, and a measure of "cognitive inhibition," has found that subjects high on inhibitory ability have longer future time "conceptions" (Levine, Spivack, Fuschillo, & Tavernier, 1959). But Mischel (1961c), using the direct measure employed in the present study and Trinidadian children, found that preference for DelR was related not to length, but to *accuracy* of time conceptions (of past events). Theoretically, there are two different things one could expect. If we assume that "immediate" subjects are unable to tolerate long time delays, we would expect a linear relationship between length of time perspective and delay preference. Realistically, however, temporal distance can be overestimated as well as underestimated. Therefore, assuming that "delay" subjects have greater reality contact, we would expect a curvilinear relationship between length of time perspective and delay preference, with "immediate" subjects making estimates that are either very short or very long, in terms of plausible realism.

METHOD

Subjects

Subjects were 162 elementary school children, from predominantly lower-middle and lower-class urban backgrounds in the Boston area. The 126 subjects (68 boys, 58 girls) of the main group (Grades 1 through 6) were from one school and were all tested on one day. In addition, 36 subjects (21 boys and 15 girls, all kindergarten children from a second, similar school in the same community), were tested separately on another day.

Main Group

Design. The purpose of the experiment called for a 6 × 5 design, in which age was varied by using six grades and length of delay interval was varied by using five different delay periods (1 day, 5 days, 1 week, 2 weeks, 4 weeks). On the basis of the school list, each class was randomly divided into five groups and each group assigned to one of the delay periods. Due to absenteeism and to the fact that the younger grades had fewer children, some of the cells had very small frequencies. Subjects from Grades 3 through 6 and Grades 1 through 2 were tested together, though subjects from each delay length category were tested separately. Thus for the purpose of administering the procedure there were 10 groups and these were assigned at random to two experimenters (*E*s). However, since the groups made up of the older children (Grades 3 through 6) were larger, one *E* tested more subjects than the other (74 vs. 52). The possible effects of this will be discussed later.

Procedure and measures. The reinforcements, chosen on the basis of pretesting and previous experience, were: a $.05 candy bar (Hershey, with almonds) today (ImR) *or* a $.10 candy bar of the same kind later (DelR, e.g., tomorrow or next week, etc.). This real choice was offered, as a reward for cooperating, at the end of the total procedure. Both candies were shown and at the promised time the subject was given the *one* he actually chose.

In order to obtain a measure of *future time perspective* the following questions were asked:

1. What year do you think the first man will reach the moon?
2. What year do you think somebody will run the mile in 3.5 minutes?
3. What year will everyone have their own airplanes, like today they have their own cars?
4. What year will somebody hit 60 home runs again?
5. How old will you be when you are grown up?
6. How old will you be when you get married?
7. How old will you be when you make the most money?
8. When will the summer vacation start?

Since their distributions showed very little spread, Questions 5 and 8 were omitted from the analysis. The answers to the other six questions were divided into thirds and subjects assigned scores from 1 (shortest) to 3 (longest) on each question; these were then summed to get a total length of time perspective score.[2]

Additional Group

Procedure. The kindergarten subjects were included in order to extend the age range downward

[2] For the younger children (Grades 1 and 2) only Questions 1 and 6 were included in the procedure, as a first step in pretesting appropriate items for this age range. So the analysis of time perspective is based on the older subjects only. Brief simple measures of acquiescence and occupational preference were also included in the procedure for the older subjects. However, since these were exploratory and are not relevant to the main purpose of the study, they will not be discussed in the present report.

as far as possible. The only procedure here was to ask subjects to print their first name on a piece of paper. They were told to indicate their choice of reward ($.05 candy today, or $.10 candy next week) by writing T (today) or L (later) on their paper, the writing being first demonstrated on the blackboard. Subjects were tested in two groups by one experimenter (E_1).

RESULTS

The data were examined for sex differences, which were not significant ($\chi^2 = 2.306$, $df = 1$, $p > .10$); therefore all subsequent analyses combine results from boys and girls.

The results on the two main dimensions, age and length of delay interval, are shown in Table 1. There is a clearly visible trend for the proportion of DelR choices to increase with age, the most marked change occurring between Grades 3 and 4. The results from the kindergarten subjects (mean age = 5.7) are consistent with those of the main group: the ratio of immediate choices to delayed is 26:10. If we consider only the age variable we have a set of proportions in seven ordered categories (Grades 1 through 6 and kindergarten). Hence a statistic that enables us to take out for trend is appropriate (Armitage, 1955). The chi square for linear regression is 23.75 ($df = 1$, $p < .001$). The nonlinear effect is not significant ($\chi^2 = 9.04$, $df = 5$, $p > .10$). Alternatively, we may do a correlational analysis, in which chronological age is one variable and choice of DelR the other. This yields a point biserial r of .65 ($N = 160$, $p < .0001$). The 95% confidence limits for this correlation coefficient are .55 and .73.

The mean IQ of children choosing DelR is 105.7 ($N = 64$, $SD = 11.25$), and of children choosing ImR 99.0 ($N = 58$, $SD = 10.78$).[3] This difference is highly significant ($CR = 3.37$, $p < .0005$). Again, we may perform an alternative analysis in terms of correlation: the point biserial r between IQ and DelR choice is .29 ($N = 122$, $p < .0005$). The 95% confidence limits for this correlation coefficient are .12 and .45. In order to test whether the correlation between preference for DelR and intelligence may be explained in terms of the correlation between preference for DelR and age (or vice versa) the correlation between age and IQ for this sample was computed. Since this was not significant ($r = .14$, $N = 122$, $.10 > p > .05$), it may be assumed that the two correlations are independent.

[3] The IQ scores are based on the Lorge-Thorndike Form A for the first and second grade subjects, and on the Pintner General Ability Test, Verbal Series, Forms A or B, for the older subjects. All scores were taken from the school records.

TABLE 1

THE RELATIONSHIP OF AGE AND DELAY INTERVAL LENGTH TO PREFERENCE FOR DELAYED VS. IMMEDIATE REWARD[a]

Delay interval	Grades[b]						Totals[c]
	1 (6.6)	2 (7.6)	3 (8.8)	4 (9.5)	5 (10.6)	6 (11.6)	
1 day	3:0	2:3	1:1	0:4	0:5	0:5	6:18
5 days	4:1	1:1	4:0	1:4	2:3	2:3	14:12
1 week	2:2	2:1	4:1	1:4	2:3	2:3	13:14
2 weeks	3:1	3:1	2:2	1:3	0:3	0:5	9:15
4 weeks	2:0	2:2	4:0	4:2	3:1	2:3	17:8
Totals[d]	14:4	10:8	15:4	7:17	7:15	6:19	59:67

[a] In each cell the first figure gives the number of immediate, smaller (ImR) choices, the second the number of delayed, larger (DelR) choices.
[b] The figures in parentheses are mean chronological ages.
[c] χ^2 for linear trend across delay intervals = 4.91, $df = 1$, $p < .05$.
[d] χ^2 for linear trend across age groups = 23.75, $df = 1$, $p < .001$.

To test the third hypothesis the total proportions in the five delay length categories (last column of Table 1) were analyzed in the same way as the proportions in the age categories. The chi square for linear regression is 4.91 ($df = 1$, $p < .05$). The nonlinear effect is again not significant ($\chi^2 = 5.35$, $df = 3$, $p > .10$). Thus the hypothesis is confirmed: the proportion of ImR choices increases linearly with length of delay interval.

In order to test for interaction between age and length of delay interval, the analysis of trend due to delay length was also carried out separately for younger and older children. If we consider only the subjects from the first to the third grade (i.e., the totals of the first three columns in Table 1) the chi square for linear regression is 0.23 and that for curvilinearity is 1.69, neither of which is significant. In the subjects drawn from the fourth to the sixth grade (i.e., the totals of the last three columns of Table 1) the chi square for linear regression is very significant ($\chi^2 = 7.08$, $df = 1$, $p < .01$) and the nonlinear ef-

TABLE 2

THE RELATIONSHIP OF FUTURE TIME PERSPECTIVE TO PREFERENCE FOR DELAYED REWARD

Condition	Time perspective[a]			
	Short	Medium	Long	
ImR	16	5	9	$\chi^2=11.72, p<.01$
DelR	9	20	22	

[a] Categorization of the distribution into closest approximation of thirds.

fect is slightly significant ($\chi^2 = 8.66$, $df = 3$, $p < .05$). Thus it would appear that the overall trend in all groups combined is attributable to the older group only, and that the younger children make no differentiation between the various time periods. However, the interpretation of this interaction is not unambiguous in view of the experimenter differences to be discussed below.

The data on length of time perspective into the future, using Grades 3 through 6 only, yield a mean of 11.27 ($SD = 2.85$) for the group choosing the immediate, smaller reward ($N = 30$); and a mean of 12.20 ($SD = 1.92$) for the group choosing the delayed larger reward ($N = 51$). Thus time perspective is slightly, though nonsignificantly, shorter in the ImR group ($t = 1.57$, $p > .10$), but significantly more variable ($F = 2.21$, $p < .02$).

In order to examine more closely the nature of the difference in variability, the distribution of the total time perspective scores was divided into thirds and thus compared with ImR vs. DelR choice. This is shown in Table 2, where it can be seen that there is a significant tendency for the ImR subjects to have either very low or very high time scores. In terms of individual questions this means, e.g., that they were more likely to give answers such as "this July" or "in the year 30,000." There are no significant differences

TABLE 3

PREFERENCE FOR DELAYED VS. IMMEDIATE REWARD IN SUBJECTS TESTED BY THE TWO EXPERIMENTERS

Condition	E_1	E_2	
ImR	17	42	$\chi^2 = 6.0$, $p < .02$
DelR	35	32	

TABLE 4

RELATIONSHIP OF DELAY PREFERENCE TO AGE, FOR EACH E SEPARATELY

Condition	E_1			E_2		
	Grades			Grades		
	1 and 2	5 and 6		1 and 2	5 and 6	
ImR	13	0	$\chi^2=15.01$, $p<.001$	11	13	$\chi^2=4.41$, $p<.02$
DelR	7	18		5	26	

in the time perspective scores for the different age groups. The correlation of time perspective length with IQ is .17 ($N = 76$, $p < .07$) which is suggestive but not strong.

Differences between Es

Table 3 shows that there is a significant difference between Es: E_1 elicits more DelR choices, E_2 elicits more ImR choices. If both experimenters tested the same numbers and kinds of subjects this difference would not affect the interpretation of the results. However, this was not possible both because of sample shrinkage and because the procedure called for first and second grades to be tested separately from the older children. Specifically, each E tested five groups, one at each delay interval—E_1 tested three younger groups and two older, E_2 tested three older groups and two younger. In addition, the first and second grades were smaller than the others. The result was that E_2 tested more older children than E_1, and also more children at the longer delay periods. In order to determine to what extent differences in Es affected the previously found relationships,

TABLE 5

RELATIONSHIP OF DELAY PREFERENCE TO DELAY LENGTH, FOR EACH E SEPARATELY

Condition	E_1			E_2		
	1 day	4 weeks		1 day	4 weeks	
ImR	1	4	$\chi^2=5.96$, $p<.02$	5	13	$\chi^2=0.02$, *ns*
DelR	15	2		3	6	

these were reanalyzed for each *E* separately.

Table 4 shows that if preference for DelR in the two youngest and two oldest age groups is compared, the results are significant for both *E*s. Moreover it should be noted that the difference in experimenters would go against the hypothesis, i.e., reduce the strength of the relationship between age and preference for DelR, since E_2, who elicited more ImR choices, also tested more subjects in the older age groups.

Table 5 shows that if preference for DelR is compared in the groups tested under the shortest (1 day) and longest (4 weeks) delay periods, the difference is significant for E_1 but not for E_2. It may be concluded that the tendency of E_2 to elicit more ImR choices masked the effect of delay length differences in subjects tested by him, and that the overall trend must be regarded as a joint function of differences in length of delay interval and reaction to the *E*.

Finally, the previously discussed interaction between age and delay interval is absent for subjects tested by E_1, and nonsignificant although in the predicted direction for subjects tested by E_2. On the basis of the present data then, the previous conclusion that for the younger children only the "now-later" contrast is relevant must be regarded as tentative. The interaction between age and delay length appears in the data, but is not unambiguously interpretable because it is confounded with *E* differences. We may state the conclusion to be drawn as follows:

1. There is a general tendency for the proportion of DelR choices to decrease with increasing length of the delay interval.

2. This tendency is especially marked in the older subjects, but this may be due to either of two possibilities: (*a*) older subjects make more sensitive temporal discriminations, *or* (*b*) a general set to make more ImR choices to the experimenter (E_2) who tested most of the older subjects.

DISCUSSION

As predicted, a strong linear relationship of preference for DelR with age, and a less strong, but still very significant relationship with intelligence, were found. Furthermore, the predicted inverse relationship of delay preference and delay interval was also evident, although it was at least partially masked by differences in the reactions to the two experimenters. It should be noted that the latter finding is in accord with those of Mahrer (1956) and Terrell (1959) who found that the experimental manipulation of incentive values in young children is extremely sensitive to differences in the behavior of the *E*.

The inverse relationship of delay preference and length of delay interval was predicted on the basis of the assumption that the expectancy of obtaining a reinforcement decreases with its temporal distance. However, the present experiment does not rule out the alternative explanation that the negative value of waiting subtracts from the positive value of the incentive, regardless of expectancy. The experiment by Mahrer (1956) showed that different levels of expectancy affect preference for delayed rewards even when delay length is held constant. The converse experiment, varying the length of the delay interval while holding expectancy constant, has yet to be done. The results of the present experiment may be regarded as a preliminary control for such a study: they show that length of delay interval affects DelR preference. But this may reflect differences in expectancy only, if the passage of time per se is unimportant, as Rotter (1954, p. 161) has suggested; or it may reflect differences in expectancy *and* differences in the aversiveness of waiting.

In the present experiment the major changes occurred at the age of 8.5–9, between the third and fourth grades. The change from a preponderance of ImR choices to a preponderance of DelR choices came at this point, and there was some suggestive evidence that the effect of delay intervals of differing length was significant above this age only. It is possible to speculate that for younger children the different lengths of time are meaningless and only the now-later contrast is relevant. The present data lend some support to such a notion, though it is doubtful whether the precise age at which the change occurs is a general "critical age." For, in the first place, different pairs of reinforcements would probably yield different cutting-points, even

though the overall trend might remain the same. Melikian (1959), for example, found that using his particular monetary reinforcements (10 fils vs. 20 fils) in his group of refugee children, the crucial age was 6, and concluded that after 6, no more increase in preference for delay is to be expected, since refugees eventually lose hope. In the present study, change occurred after 6, but none after 9. It would seem more plausible to interpret such results as reflecting the fact that the differences in the particular reinforcements chosen cease to be very important after a certain age. Naturally, this age will vary with the values chosen. One might expect that the "critical age" could be located at any desired point on the age scale by appropriate manipulation of reinforcement values. The problem in demonstrating an extended age trend is one of finding the appropriate *measure*, rather than one of locating specific stages in development. Secondly, the time perspective scores show no significant change at 8 or 9, as one would expect if this were indeed the general critical age.

The results on time perspective tend to support the second of the two possibilities mentioned earlier, viz., that delay subjects make moderate, realistic estimates of future events, whereas immediate subjects make either extremely short or extremely long estimates. There is also a slight tendency for the immediate subjects to have shorter time perspective. Also, the tendency for the more intelligent subjects to have longer time perspectives confirms the findings of Levine, Spivack, Fuschillo, and Tavernier (1959) on adolescents. That these results may be further complicated by a "ceiling effect" in the delay measure is suggested by some recent findings (Mischel, Gershuni, & Vollmar, 1960). Using a continuous rather than a dichotomous delay preference measure and a measure of length of time span of TAT stories, it was found that subjects with the highest and lowest DelR scores, i.e., those located at the extremes of the distribution of DelR scores, have shorter time spans than those with medium DelR scores. This effect would blur differences in time perspective (assuming the two time measures are roughly analogous) when a crude measure of delay preference such as that of the present study is used. For in the DelR group there would be included both subjects with long time scores (the moderate delayers) and subjects with short time scores (the extreme delayers).

SUMMARY

Using a simple choice measure of preference for delayed reward on elementary school children in the Boston area, ranging in age from 5 to 12, and under five different lengths of delay interval, the following hypotheses were tested and confirmed: (*a*) preference for delayed reward is positively related to age; (*b*) positively to intelligence; (*c*) and negatively to length of the delay interval. The findings on Hypothesis *c* were not entirely unambiguous owing to the extreme sensitivity of the delay measure to differences in the *E*. An additional finding was that subjects preferring the immediate reward tend to have more variable future time perspectives and that length of time perspective is slightly related to IQ, but not to age.

REFERENCES

ARMITAGE, P. Tests for linear trends in proportions and frequencies. *Biometrics*, 1955, **11**, 375–386.

FREUD, S. Formulations regarding the two principles in mental functioning. (Originally published 1911) In, *Collected papers*. Vol. IV. New York: Basic Books, 1959. Pp. 13–21.

LEVINE, M., SPIVACK, G., FUSCHILLO, JEAN, & TAVERNIER, ANN. Intelligence, and measures of inhibition and time-sense. *J. clin. Psychol.*, 1959, **15**, 224–226.

LEVINE, M., SPIVACK, G., & WIGHT, B. The inhibition process, Rorschach human movement responses, and intelligence: Some further data. *J. consult. Psychol.*, 1959, **23**, 306–313.

MAHRER, A. The role of expectancy in delayed reinforcement. *J. exp. Psychol.*, 1956, **52**, 101–105.

MELIKIAN, L. Preference for delayed reinforcement: An experimental study among Palestinian Arab refugee children. *J. soc. Psychol.*, 1959, **50**, 81–86.

MISCHEL, W. Preference for delayed reinforcement: An experimental study of a cultural observation. *J. abnorm. soc. Psychol.*, 1958, **56**, 57–61.

MISCHEL, W. Delay of gratification, need for achievement, and acquiescence in another culture. *J. abnorm. soc. Psychol.*, 1961, **62**, 543–552. (a)

MISCHEL, W. Father-absence and delay of gratification: Cross-cultural comparisons. *J. abnorm. soc. Psychol.*, 1961, **63**, 116–124. (b)

MISCHEL, W. Preference for delayed reinforcement and social responsibility. *J. abnorm. soc. Psychol.*, 1961, **62**, 1–70. (c)

MISCHEL, W., GERSHUNI, C., & VOLLMAR, J. Studies on delay of gratification and time sense. Unpublished manuscript, Harvard University, 1960.

RAPAPORT, D. On the psychoanalytic theory of thinking. *Int. J. Psychoanal.*, 1950, **31**, 161–170.

ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.

SINGER, J. L. Delayed gratification and ego-development: Implications for clinical and experimental research. *J. consult. Psychol.*, 1955, **19**, 259–266.

SINGER, J. L., WILENSKY, H., & MCGRAVEN, VIVIAN G. Delaying capacity, fantasy, and planning ability: A factorial study of some basic ego functions. *J. consult. Psychol.*, 1956, **20**, 375–383.

SPIVACK, G., LEVINE, M., FUSCHILLO, JEAN, & TAVERNIER, ANN. Rorschach movement responses and inhibition processes in adolescents. *J. proj. Tech.*, 1959, **23**, 462–466.

TERRELL, G. Manipulatory motivation in children. *J. comp. physiol. Psychol.*, 1959, **52**, 705–709.

(Received April 6, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 432–437

ANXIETY AND AVOIDANCE BEHAVIOR IN SCHIZOPHRENICS IN RESPONSE TO PARENTAL FIGURES [1]

JAMES C. BAXTER [2] AND JOSEPH BECKER [3]

John Umstead Hospital, Butner, North Carolina

The cardinal importance of deviant familial relationships in the background of schizophrenic patients has long been recognized by personality theorists. A sizable literature in the area strongly suggests that the schizophrenic comes from a family milieu differing in systematic ways from that of his nonschizophrenic peer (Bateson, Jackson, Haley, & Weakland, 1956; Bowen, 1959; Lidz & Fleck, 1960; Wynne, Ryckoff, Day, & Hirsch, 1958). Recent reports have also drawn attention to differences in patterns of family relationships which can be distinguished between subgroups of schizophrenics. By differentiating patients along a continuum of premorbid personality adequacy, as defined by the Phillips (1953) Scale of Premorbid Adjustment, Rodnick and Garmezy (1957; Garmezy & Rodnick, 1959) and their associates have identified reliable patterns of asymmetrical authority and submission, as well as differences in interpersonal conflict in the family. According to these studies, male schizophrenics who have attained a relatively high degree of psychosocial maturity and sexual participation prior to their psychosis (Goods) tend to come from homes dominated by a tyrannical father, with the mother occupying a comparatively weak and ineffectual role. On the other hand, patients who have failed to establish stable sexual relationships and who have been relative social isolates prior to the appearance of the disorder (Poors) tend to come from mother dominated homes high in conflict and interpersonal aggression (Farina, 1960; Garmezy, Farina, & Rodnick, 1960; Garmezy & Rodnick, 1959; Garmezy, Stockner, & Clarke, 1959; Rodnick & Garmezy, 1957). The premorbid adequacy variable has also been found to bear a significant relationship to differences in reaction sensitivity to social censure, which, in turn, has been shown to relate to performance deficits on a number of experimental variables. This line of investigation has suggested that a highly generalized avoidance response is more prepotent in Poors than in Goods, as a function of increased sensitivity to censure, and that this response tendency functions as an interference to integrated problem solving behavior. Rodnick and Garmezy (1957) have proposed a relationship between the differences in familial backgrounds in these groups and their differential sensitivity to censure and rebuff (p. 146).

The present study was aimed at evaluating several implications of this position in terms of anxiety and avoidance behavior aroused in schizophrenics in response to parental figures. If exposure to differing patterns of parental dominance and conflict has a significant effect in producing the reaction sensitization determining the behavior deficits observed, then differences in anxiety and avoidance should emerge in thematic productions developed in response to pictorial figures symbolizing these early traumatic (censorious) interactions. Accordingly, the following predictions were made: (*a*) Poor premorbid schizophrenics should produce more anxiety related thematic imagery in response to the maternal figure than to the paternal figure, while Goods should show the opposite effect; (*b*) Poors should produce a significantly greater amount of avoidance related imagery in response to both parental figures; and (*c*) Poors should produce more avoidance in response to the mother than the father, with Goods showing the reverse pattern.

METHOD

Subjects. The sample was chosen from a pool of 66 male patients undergoing treatment at the Veterans Administration Hospital, Roanoke, between 1949 and 1952, who had been given the diagnosis of schizophrenia by qualified staff psychiatrists, and

[1] This research was supported in part by National Institute of Mental Health Grant OM-124, "Family Centered Treatment in Schizophrenia."

[2] Now at the University of Kentucky.

[3] Now at the University of Illinois.

who had given responses to cards 6BM (mother-son) and 7BM (father-son) of the TAT during routine, individually administered psychological testing.[4] The length of time in the hospital before testing was less than 30 days in all cases. Only subjects between 20 and 40, who had been reared by an intact family until the age of 18 were included.[5] Patients whose diagnosis was complicated by indications of mental deficiency or organic brain pathology or who had received radical somatic therapy within 90 days before testing or were currently undergoing regular drug therapy were excluded. Some patients occasionally taking small amounts of sedatives to facilitate sleep were included, if the most recent administration was no later than the night before testing. The number of such patients included was small, and no difference in the proportion between groups was noted.

Thirty schizophrenics fulfilled all criteria imposed. Premorbid personality ratings were obtained from the case history of each patient by rating the degree and adequacy of social participation and sexual behavior prior to the onset of psychosis on the Phillips (1953) Scale of Premorbid Adjustment. This scale differentiates patients along a 30-point continuum of premorbid adequacy, with higher scores reflecting greater degrees of social isolation and sexual disturbance. The experimental groups were organized by taking contrasting extremes on this dimension. Patients with a score of 15 or less were assigned to the group of Goods ($N = 12$), while patients obtaining scores greater than 15 were classified as Poors ($N = 18$). The ranges of the ratings obtained were 3–15 and 17–27 for Goods and Poors, respectively. Mean age and years of education for both groups are presented in Table 1. Inspection of these data reveals that the groups are quite comparable on these variables, with no significant differences appearing. Less precise data also suggested comparability between groups for diagnostic subclassification of schizophrenia and socioeconomic position.

Scoring procedure. Only responses to Cards 6BM and 7BM of the TAT were considered in the analysis. These cards were chosen for study because they typically elicit thematic imagery relevant to parent-child relationships, and they are most frequently interpreted directly as mother-son and father-son interactions, respectively. A rating schedule was devised which would reflect various component dimensions of anxiety and avoidance behavior, and would permit reliable judgments to be made from oftentimes brief, elliptically recorded TAT protocols. Both stylistic and contentual aspects of the protocols were considered relevant to the variables under study, and the weightings given the several dimensions chosen roughly corresponded to the experimental and clinical validity attributed to them.

TABLE 1

MEAN AGE AND EDUCATIONAL LEVEL OF GOOD AND POOR PREMORBID GROUPS

Level	*N*	*SD*	*M*
Age			
Goods	12	4.89	27.9
Poors	18	4.26	27.4
Education			
Goods	12	2.86	10.2
Poors	18	2.79	10.4

A number of style and content variables in the linguistic productions of a variety of populations have been shown to be significantly related to manifest anxiety (e.g., Dollard & Mowrer, 1947; Mahl, 1959a; Meadow, Greenblatt, & Solomon, 1953; Osgood, 1957; Phillips & Smith, 1953; Walker & Atkinson, 1958). Five areas which seem to have achieved most extensive validation and clinical utility were chosen for inclusion in an anxiety rating schedule. The areas chosen were: indications of verbal interferences, viz., repetitions, sentence incompletions, sentence corrections, tongue slips, and stuttering, as defined by Mahl (1959b); statements of inability to comply with the task; signs of disorganization of the theme, e.g., digressions, irrelevancies, "private" associative connections, etc.; content of a predominately negative affective tone; and the development of a bleak, pessimistic outcome. Variable weights assigned to components of each of the areas were combined into a 12-point anxiety index, with higher scores indicating greater anxiety. A comparable index of avoidance related behavior was developed by translating the results of several divergent studies into scales applicable to TAT protocols (e.g., Bindra, 1959; Miller, 1959; Scott, 1958). Four dimensions were chosen and assigned varying weights. The general categories were: the amount of displacement of person, place, and/or time evident in the theme; the predominant direction of movement implied, i.e., toward or away from an affectively meaningful, spatially proximal relationship; the amount of verbal productivity in thought units (Dollard & Mowrer, 1947); and the degree to which objectively present figures were ignored in the themes produced. A 12-point avoidance index was obtained by summing component scores of the several categories, with higher values indicating a greater degree of avoidance.

All ratings were made by one judge from typescripts of the original protocols, coded to prevent identification of group membership. Two additional

[4] We are indebted to Lon Ussery, University of North Carolina, for making the protocols and necessary information available to us and for doing the ratings of premorbid adjustment. Burk Smith, Veterans Administration Hospital, Roanoke, Virginia, also provided invaluable assistance in obtaining drug and somatic therapy information on the patients.

[5] An intact family was defined by the absence of social history information indicating death, separation, or divorce of the parents before the patient had reached age 18. Fifty percent of the Goods and 18% of the Poors who were otherwise eligible were discarded on the basis of this criterion.

TABLE 2

ANALYSIS OF THE VARIANCE OF ANXIETY INDEX SCORES

Source	*SS*	*df*	*MS*	*F*
Between-patients	166.48	29		
Premorbid level	2.84	1	2.84	<1.00
Error	163.64	28	5.84	
Within-patients	156.50	30		
Parental figure	12.15	1	12.15	2.64
Premorbid × Parent	15.21	1	15.21	3.30**
Error	129.14	28	4.61	
Total	322.98	59		

** Significant at the .05 point, one-tailed.

raters also independently scored 15 randomly chosen protocols in order to estimate interrater reliability.[6] Product-moment correlations between raters were found to vary from .70 to .88 for the anxiety index, and from .93 to .97 for the avoidance index, with mean correlations of .78 and .95, respectively. The reliability of the anxiety scores was somewhat lower than anticipated, although it was considered adequate for present purposes, in view of the protean nature of content expressions of anxiety and the variabilty in quality of recording found in the TATs. The reliabilty of the avoidance measure, on the other hand, was quite adequate.

RESULTS

The first hypothesis predicted that the pattern of anxiety related imagery would differ within the responses of Goods and Poors, with Poors expressing more anxiety in response to the mother figure and Goods more in response to the father figure. Accordingly, the anxiety scores for both groups were entered in a Type I mixed analysis of variance table (Lindquist, 1953). The assumption of homogeneity of score variance within cells was evaluated by Hartley's F_{max} test (Edwards, 1954) and found to be tenable ($p > .05$). A summary of the analysis of anxiety scores is presented in Table 2. It can be seen that the predicted interaction of anxiety reactivity and parental figure is significant at the .05 point of confidence.[7] It is also apparent that the over-all level of anxiety between groups and mean responsivity to the mother as opposed to the father figure do not differ significantly. The mean anxiety scores for the groups are represented graphically in Figure 1. The curves of anxiety responsivity interact as predicted, with Poors reflecting greater responsivity to the mother than the father figure, while the pattern for Goods is in the opposite direction.

The avoidance scores permitted evaluation of the remaining hypotheses. That is, the over-all level of avoidance was expected to be greater for Poors, and the patterns of avoidance behavior were expected to differ between Goods and Poors. Scores were arrayed in a Type I table, and tested for homogeneity of cell variances. Hartley's F_{max} test indicated that such an assumption was questionable ($.10 > p > .05$), so that a square root transformation of all scores was done in order to remove heterogeneity within cells. These transformed scores were then analyzed for group and pattern differences. Table 3 summarizes the results of these comparisons. It can be seen that a significant difference in the level of avoidance behaviors exists between groups ($p < .025$), with Poors exceeding Goods in response to both figures. The expected difference in the patterning of avoidance behavior did not occur, however ($p > .05$). Inspection of individual means indicated that, while Goods tended to show the expected increase in avoidance responsivity to the father figure, Poors were equally avoidant in response to both parental figures.

It is possible that the failure of discriminative avoidance could be accounted for as a result of the excessively high level of avoidance behavior in the Poors. It could be ar-

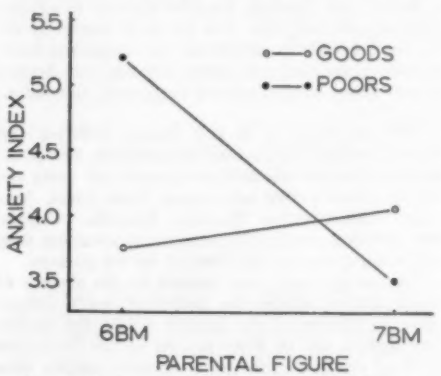

FIG. 1. Anxiety responsivity to parental figures.

[6] We are indebted to Walter Hooks, John Umstead Hospital, for assisting the authors in making the ratings.

[7] One-tailed tests of significance are reported for the hypotheses, since directional predictions were made in each case.

gued that the presence of a highly prepotent avoidance tendency vitiated discriminative response patterns in this group, and thus led to grossly overgeneralized avoidance behavior in the Poors. Garmezy (1952; Rodnick & Garmezy, 1957) has found evidence to support such a position when performance is observed under conditions of experimentally imposed social censure. In an attempt to evaluate the possibility that patterned avoidance was precluded as a result of the prepotence of avoidance behavior in the Poors in the present situation, matched subgroups of Goods and Poors were selected on the basis of total avoidance behavior. Nine individually matched pairs of Goods and Poors were found, each of which differed less than .2 point on total avoidance. Discriminative avoidance scores were computed for these 18 subjects by dividing avoidance scores on Card 6BM by the corresponding scores on 7BM. Thus, scores less than unity indicated greater avoidance in response to the father card, while scores greater than unity reflected more avoidance to the mother figure. The Mann-Whitney U test was applied to the groups, and found to be nonsignificant ($U = 37.0$, $p > .05$), although the rank difference was in the predicted direction. Accordingly, no evidence of consistent patterning of avoidance behavior was found, even when total avoidance was controlled. A further analysis of these matched subgroups did prove suggestive, however. The comparison of group variability of the discriminative avoidance scores indicated considerable heterogeneity between groups, with the Goods significantly more variable than the Poors ($F = 3.87$, $df = 8/8$, $p < .05$). The ranges of these scores were .42–1.65 for the Goods, and .77–1.29 for the Poors. Since other studies have consistently found Goods to be more uniform than Poors in their performance on a number of tasks (Garmezy & Rodnick, 1959; Rodnick & Garmezy, 1957), this result is unexpected and suggests that discriminative avoidance behavior may have occurred in the Goods, although without the characteristic pattern predicted.

TABLE 3

ANALYSIS OF THE VARIANCE OF AVOIDANCE INDEX SCORES

Source	*SS*	*df*	*MS*	*F*
Between-patients	7.78	29		
Premorbid level	1.07	1	1.07	4.46*
Error	6.71	28	.24	
Within-patients	4.32	30		
Parental figure	.10	1	.10	<1.00
Premorbid × Parent	.19	1	.19	1.36
Error	4.03	28	.14	
Total	12.10	59		

* Significant at the .025 point, one-tailed.

DISCUSSION

The results of the present study demonstrate a significant difference in the pattern of anxiety related behavior aroused in Good and Poor premorbid schizophrenics in response to pictorial representations of parental figures. The differential cue properties of these figures conform to expectations arising from the contrasting familial backgrounds to which the groups have presumably been exposed. Therefore, substantial support is found for the inferred asymmetry in parental dominance and submission in the backgrounds of these groups, with Goods coming from father dominant homes and Poors from a mother dominated background. It is also of interest to note the absence of group differences in total anxiety related behavior expressed. Under the present conditions, the groups cannot be distinguished on the basis of this over-all effect.

A reliable group difference in the level of avoidance related behavior was also found to characterize these patients. That is, under essentially comparable levels of total anxiety, Poors were significantly more avoidant in response to both parental figures. This result is in agreement with other observations pointing to the increased prepotence of avoidance behavior in patients attaining Poor premorbid adjustments, and thus lends further support to the position advanced by Rodnick and Garmezy (1957). The present data further demonstrate the utility of distinguishing patients along a continuum of premorbid personality adequacy. The efficacy of distinguishing subgroups of schizophrenics, apart from diagnostic subtype, has been debated for some years (Bellak, 1958; Garmezy & Rodnick, 1959). The present data concerning the patterning of anxiety and the differential pre-

potence of avoidance tendencies support the contention that such distinctions are warranted in terms of identifying subpatterns of response characteristics.

The results of the third hypothesis are less clear. In failing to confirm the predicted difference in the patterning of avoidance behavior, the data fail to agree with predictions drawn from the Rodnick-Garmezy position. Several limitations must be considered in qualifying the interpretation of this result, however. First, it is possible that the high level (prepotence) of avoidance behavior in the Poors interfered with discriminative avoidance responding, and hence concealed patterns which may be present under lower levels of response. Subgroups of Goods and Poors were selected to partially control for this possibility by equating total avoidance behavior. Comparison of discriminative avoidance scores in these matched groups also failed to reveal the differences predicted. Groups equated in this way necessarily match patients within an intermediate range of avoidance, however, and do not provide an unequivocal comparison. More adequate controls would be provided by patients on tranquilizing drugs, or patients who have achieved clinical remission of schizophrenic symptoms. A further difficulty is the order of presentation of the thematic cards. Direct evidence on card order is not available, although convention would lead one to expect that Card 6BM preceded 7BM in most cases. If recovery from threatening (censorious) stimuli is delayed in schizophrenics, the patterned (discriminated) response in Poors would be reduced, since this group is assumed to be more avoidant of the maternal figure, i.e., Card 6BM, given first. In view of these limitations in experimental controls, clear interpretation of the failure of the third hypothesis cannot be made from the present results. Evidence was found, however, to suggest that Goods are more discriminative in their avoidance behavior, since score variability for this group was significantly greater than for Poors. This result must be considered tentative, in view of the small sample size and the absence of subsidiary variables capable of reducing this variance to meaningful patterns.

SUMMARY

Thematic productions in groups of Good and Poor premorbid schizophrenics were analyzed in terms of the relative amounts of anxiety and avoidance related imagery produced in response to cards representing parent-child relationships. Three hypotheses were specified concerning the patterning of these responses: Poor premorbid patients were expected to produce more anxiety related imagery in response to the mother than the father figure, while Goods were expected to show the reverse effect; Poors were expected to be more avoidant in response to both parental figures; and the pattern of avoidance behavior between groups was expected to differ.

The first two hypotheses were supported by the present data. These results were interpreted as supporting the position advanced by Rodnick and Garmezy and associates. The differential patterning of anxiety behavior in these groups is congruent with patterns of parental dominance and conflict assumed to have occurred in the developmental backgrounds of the patients. The increased level of avoidance behavior in the Poors is also in accord with the presumed increased prepotence of avoidance tendencies in this group. The prediction concerning differential patterning of avoidance behavior failed, however, to attain significance.

REFERENCES

BATESON, G., JACKSON, D., HALEY, J., & WEAKLAND, J. Toward a theory of schizophrenia. *Behav. Sci.* 1956, **1**, 251–264.

BELLAK, L. *Schizophrenia: A review of the syndrome*. New York: Logos, 1958.

BINDRA, D. *Motivation: A systematic reinterpretation*. New York: Ronald, 1959.

BOWEN, M. Family relationships in schizophrenia. In A. Auerback (Ed.), *Schizophrenia: An integrated approach*. New York: Ronald, 1959. Pp. 147–178.

DOLLARD, J., & MOWRER, O. H. Measuring tension in written documents. *J. abnorm. soc. Psychol.*, 1947, **42**, 3–32.

EDWARDS, A. L. *Statistical methods for the behavioral sciences*. New York: Rinehart, 1954.

FARINA, A. Patterns of role dominance and conflict in parents of schizophrenic patients. *J. abnorm. soc. Psychol.*, 1960, **61**, 31–38.

GARMEZY, N. Stimulus differentiation by schizophrenic and normal subjects under conditions of

reward and punishment. *J. Pers.*, 1952, **20**, 253–276.

GARMEZY, N., FARINA, A., & RODNICK, E. H. The structured situational test: A method for studying family interaction in schizophrenia. *Amer. J. Orthopsychiat.*, 1960, **30**, 445–452.

GARMEZY, N., & RODNICK, E. H. Premorbid adjustment and performance in schizophrenics: Implications for interpreting heterogeneity in schizophrenia. *J. nerv. ment. Dis.*, 1959, **129**, 450–466.

GARMEZY, N., STOCKNER, CAROL, & CLARKE, A. R. Child-rearing attitudes of mothers and fathers as reported by schizophrenic and normal control patients. *Amer. Psychologist*, 1959, **14**, 333. (Abstract)

LIDZ, T., & FLECK, S. Schizophrenia, human integration, and the role of the family. In D. Jackson (Ed.), *The etiology of schizophrenia*. New York: Basic Books, 1960. Pp. 323–345.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

MAHL, G. F. Exploring emotional states by content analysis. In, *Trends in content analysis*. Urbana: Univer. Illinois Press, 1959. Pp. 89–130. (a)

MAHL, G. F. Measuring the patient's anxiety during interviews from "expressive" aspects of speech. *Trans. N. Y. Acad. Sci.*, 1959, **21**, 249–257. (b)

MEADOW, A., GREENBLATT, M., & SOLOMON, H. C. *Frontal lobes and schizophrenia*. New York: Springfield, 1953.

MILLER, N. E. Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation, and social learning. In S. Koch (Ed.), *Psychology: A study of a science*. New York: McGraw-Hill, 1959. Pp. 196–292.

OSGOOD, C. E. Motivational dynamics of language behavior. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 348–423.

PHILLIPS, L. Case history data and prognosis in schizophrenia. *J. nerv. ment. Dis.*, 1953, **117**, 515–525.

PHILLIPS, L., & SMITH, J. G. *Rorschach interpretation: Advanced technique*. New York: Grune & Stratton, 1953.

RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 109–183.

SCOTT, W. A. The avoidance of threatening material in imaginative behavior. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society*. Princeton: Van Nostrand, 1958. Pp. 572–585.

WALKER, E. L., & ATKINSON, J. W. The expression of fear-related motivation in thematic apperception as a function of proximity to an atomic explosion. In J. W. Atkinson (Ed.), *Motives in fantasy, action, and society*. Princeton: Van Nostrand, 1958. Pp. 143–159.

WYNNE, L. C., RYCKOFF, I. M., DAY, JULIANA, & HIRSCH, S. J. Pseudo-mutuality in the family relations of schizophrenics. *Psychiatry*, 1958, **21**, 205–220.

(Received April 10, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 438–445

RESPONSE TO CONFLICT AS DETERMINED BY PRIOR EXPOSURE TO CONFLICT [1]

LEONARD WORELL

Oklahoma State University

Clinical descriptions frequently contain statements indicating that people who are faced with strong conflicts demonstrate inefficient behavior not only in relation to these strong conflicts but also in resolving comparatively mild conflicting situations. A "neurotic" individual, for example, who presumably reacts with strong conflicting tendencies only in certain areas, may be characterized as generally indecisive. Thereby, the implication is conveyed that the person's exposure to severe conflict has significantly contributed to his inefficient handling of weaker conflicts. In spite of these recurrent observations, the problem of the generalization of the effects of conflict has received scant experimental attention. With small exception research in the area of conflict has concentrated on (*a*) the behavioral consequences of variations in several characteristics of conflicting responses, such as their relative strengths, approach-avoidance, etc., and (*b*) the role of individual differences in conflict resolution (Andreas, 1958; Berlyne, 1960; Block & Peterson, 1955; Brown, 1942; Cartwright, 1941; Castaneda & Worell, 1961; Hovland & Sears, 1938; Kaufman & Miller, 1949; Scodel, Ratoosh, & Minas, 1959; Sears & Hovland, 1941; Worell & Castaneda, 1961b; Worell & Hill, in press. This paper continues (Worell & Castaneda, 1961a) the investigation of a third area; that in which conflict serves as an antecedent to performance in other than the conflict situation. The general question under consideration is as follows: Does responding to conflict have an effect on the resolution of relatively weak conflict situations?

Our attention will be restricted to two separate but related facets of this general problem. First, we will describe two opposing theoretical formulations that deal with the effects on the individual of exposure to different degrees of conflict. These hypothetical consequences will then be used to anticipate the performance of subjects in a subsequent conflict situation. Second, we will consider the issue of the degree of conflict itself. In this connection, the importance for performance of the absolute and relative strengths of conflicting tendencies will be examined.

The two theoretical approaches will be designated as the *competing response* and the *dynamogenic* hypotheses. These hypotheses may be, respectively, subsumed under the long-standing distinction made by psychologists between associative or learned and nonassociative approaches. According to the competing response view, it is assumed that responding to different levels of conflict leads to the learning of differing conflict-specific responses. For example, a person who is protractedly exposed to relatively strong conflict might be expected to learn such responses as withholding a decision or considering each alternative more carefully, etc. Then, in new but similar situations, the individual may be expected to invoke those behaviors which he has previously learned. Thus, with the competing response view, the effects of conflict are expected to have a limited generality—limited by the similarities between earlier and later conflict situations.

Sharply contrasted to the foregoing position is the dynamogenic view which proposes that conflict generates motivational consequences. The general suggestion that conflict may possess tension arousing or drive properties has been frequently made (Brown & Farber, 1951; Lewin, 1933; Miller, 1944; Miller & Stevenson, 1936). In addition, several studies designed to test this suggestion have agreed in finding heightened force of response as the degree of conflict is increased (Castaneda & Worell, 1961; Finger, 1941; Worell & Cas-

[1] The present research was assisted by a grant from the Research Foundation of Oklahoma State University.

taneda, 1961b). These results are consistent with the view that one consequence of conflict is an elevation of drive or motivation and that increasing degrees of conflict produce greater increments in drive. If conflict generates drive, then the effects of conflict on performance in a contiguously presented task would be expected to be evident irrespective of the similarity between the conflict and later performance situation. This is based on a view of drive as an indiscriminate energizer of performance (Hull, 1943).

Findings relevant to these alternatives were recently obtained by Worell and Castaneda (1961a), where the generalization of the effects of conflict was examined in relation to learning. In two experiments, subjects were exposed to either high or low experimentally induced conflict prior to the presentation of each pair of two verbal paired-associates tasks. Different degrees of conflict were aroused by the simultaneous exposure of two lines of either equal or unequal length, while the paired-associates lists contained words with either high or low amounts of intralist competition. The findings indicated that the conflict conditions did not significantly affect performance.

The foregoing findings are not compatible with expectations derived from the dynamogenic hypothesis. On the other hand, according to the competing response view, no necessary effects of conflict would be anticipated since the conflict and learning conditions were quite dissimilar. Consequently, in the present study, the degree of similarity between the conflict situation and the subsequent performance situation has been increased. Descriptively, subjects here are exposed to varying degrees of discrimination conflict and are then all placed into a weak discrimination conflict situation. Thus, the responses required in both the earlier and later conflict situations are the same, but the stimulus similarity changes to varying degrees for the subjects in the various conflict groups. In these situations, our interest is in a particular performance measure—namely, the time to initiate a decision or reaction evocation latency.

Given a weak conflict situation following differential conflict training, the following predictions would be made with the competing response and dynamogenic positions. From the standpoint of the competing response view, subjects who have been formerly trained under strong conflict should be slower in the weak conflict situation than those who had experienced less severe conflict training conditions. In this it is assumed that prior exposure to strong conflict has led to the learning of conflict-specific responses which are transferred to the new but similar situation. On the other hand, from the dynamogenic view, subjects who were formerly exposed to strong conflict should be faster than subjects exposed to weaker conflicts. This is based on two considerations: first, that increased conflict is associated with increased drive, and secondly, that increased drive, as determined by the Taylor Manifest Anxiety scale, has been found to produce faster speeds in the same relatively weak conflict situation as used here by Worell and Castaneda (1961b).

The second concern of this paper is with the problem of the degree or amount of conflict. Writings on conflict have frequently made reference to the importance of either or both the relative and absolute strengths of conflicting tendencies in defining amount of conflict. In general, studies designed to determine the effects of these two characteristics on temporal performance have varied the strengths of the conflicting tendencies by variations in either training procedures or stimulus intensity. The justification for using training procedures seems apparent; the logic for employing variations in the intensity of two or more simultaneously presented stimuli seems to be based on the repeated finding that increasing response strengths tend to be associated with increasing intensities of *single* stimuli.

The results of studies aimed at testing the assumption that the absolute strengths of competing tendencies are influential in determining the amount of conflict have been either negative or at best equivocal. For example, Berlyne (1957) exposed subjects to either two equally bright or two equally dim stimuli and found that the weaker intensities were associated with significantly *longer* response times. Castaneda and Worell (1961) found a similar, but insignificant, trend for

increased speed of response to occur with increased intensities in an experiment which presented subjects with either two relatively bright or two relatively dim lights separated by equal log brightness differences. None of these results is in accord with expectations based on absolute strengths (Andreas, 1958; Bilodeau, 1950), since increased intensities should have produced longer response times. However, Andreas has obtained results agreeing with these expectations in his finding that response times were lengthened with increased amounts of training to the conflict stimuli.[2] In contrast, Bilodeau, also employing training procedures designed to increase the strengths of the competing tendencies, failed to find a significant decrease in movement time as the absolute strengths increased.

In contrast to the findings with absolute strengths, the results of studies directed toward the effects of the absolute *difference* in strength between or the relative strengths of conflicting tendencies on temporal performance have been with few exceptions positive. Thus, studies by Bilodeau (1950), Brown (1942), Cartwright (1941), Castaneda and Worell (1961), Korman (1960), Worell and Castaneda (1961b), and Worell and Hill (1962) have agreed in finding that speed of response increases as the competing tendencies become more unlike in strength. However, no differences in temporal performance as a function of variations in relative strength have been obtained by both Andreas (1958) and Broadhurst (1957).[3]

Aside from empirical examinations of the problem of the amount of conflict, at least two theoretical formulations have been advanced which specifically incorporate either or both the relative and absolute strengths of conflicting tendencies in their approaches. First, in a detailed attempt to integrate a construct of frustration within a Hullian framework, Brown and Farber (1951) developed the position that frustration arose as a consequence of the magnitudes of *both* the absolute and relative strengths of competing tendencies. Moreover, their proposal suggests that somewhat greater importance should be given to the absolute strengths in producing frustration. In contrast, in a recent paper (Worell, 1961) where an attempt was made to account for the performance of differentially motivated subjects in learning and performance tasks involving competing response tendencies, a set of theoretical propositions based *solely* on the relative strengths of conflicting tendencies was discussed. In view of the rather general interest in the problem of the degree of conflict, the present study will examine the effects of the relative and absolute strengths of conflicting tendencies on both immediate performance and performance in a subsequently encountered conflict task.

METHOD

Design. All subjects were initially exposed to a low conflict condition in a brightness discrimination apparatus for a predetermined series of trials. In part, this procedure was used to assess whether the subjects in the subsequently separated groups demonstrated initially equivalent performances. Following exposure to low conflict, subjects were placed into either high, intermediate, or low conflict conditions for an additional fixed number of trials. The simultaneous presentation of either two equally dim or two equally bright lights constituted the high conflict condition. Intermediate conflict was determined by the exposure of a relatively bright and dim light, while low conflict consisted of the appearance of a very bright and very dim light. In all, five groups were given differential conflict training. Two groups were exposed to high conflict, two groups received intermediate conflict, and one group was given low conflict. Comparisons between the high, intermediate, and low conflict conditions provide information relevant to the effects of the relative strengths of conflicting tendencies, while comparisons between the two groups in each of the high and intermediate conditions are relevant to the role of the absolute strengths. To assess the effects of these conditions, all subjects were finally placed in the original low conflict situation.

Apparatus. The apparatus is the same as that described elsewhere (Worell & Hill, 1962). Briefly, it consisted of two windows set into a black panel, a start platform attached to the base of the panel with a visual ready signal directly above, and a push button on both the left and right edges of the start platform. A Standard Electric timer was

[2] The studies performed by Andreas (1958) and Broadhurst (1957) used measures of temporal performance that were composites of reaction evocation time and what may be termed "movement" time. Since the meaning of a composite measure may be expected to vary considerably depending on what is involved in the movement performance, it is difficult to compare the results of studies using such a measure with those using reaction evocation latency.

[3] See Footnote 2.

activated with the illumination of the windows and was terminated when the subject removed his hand from the start platform. The lights in both windows were turned off by the subject pressing either push button. Reaction latency, or time from the illumination of the windows to removal of the hand from the start platform, was recorded to the nearest 1/100 second. These reaction time scores were converted to reciprocals (1/RT × 1000), so that all performances are presented in terms of speed of response.

The stimuli consisted of four different brightnesses, 2.8, 8.8, 90.9, and 291.4 foot candles, obtained by measurements at each window with a Macbeth illuminometer. The two brightest stimuli (90.9 and 291.4) and the two dimmest (2.8 and 8.8) were separated by differences of approximately .50 log units, while the two intermediate stimuli (8.8 and 90.9) were separated by an approximate 1.0 log unit difference.

Procedure. Each subject, participating individually, was instructed to push the button below either the "brighter" or "dimmer" light, depending upon the condition to which he was assigned. The subject was informed that the experiment was concerned with the accuracy of visual discrimination. The instructions did not contain any implication that speed was considered important.

All subjects were given a total of 64 discrimination trials. The first 16 of these constituted familiarization trials with low conflict; the intermediate 24 trials presented the experimentally induced conflict conditions; and the final 24 trials presented low conflict to all subjects again. In those conditions where two unequally bright stimuli were administered, the position of the correct stimulus was randomized such that the same stimulus did not appear more than three consecutive times in the same location.

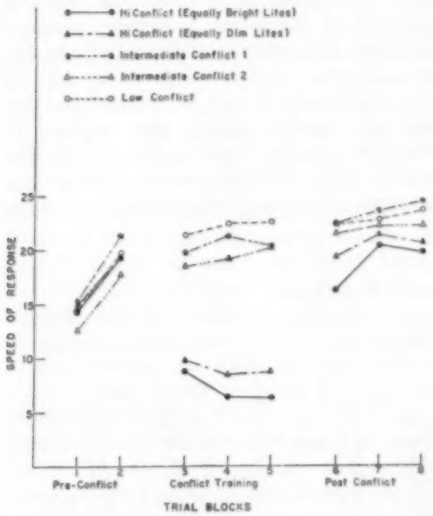

FIG. 1. Response speeds over trials for the five conflict groups during preconflict, conflict training, and postconflict conditions. (Each trial block represents the average performance on eight trials. Intermediate Conflict 1 corresponds to Group III; Intermediate Conflict 2 to Group IV [see Table 1].)

TABLE 1

MEAN SPEEDS OF RESPONSE

Group[a]	I[b]	Preconflict performance	Conflict training	Post- minus preconflict performance
I (291.4 and 291.4)	B	19.14	7.86	−3.98
	D	14.37	6.34	.43
II (2.8 and 2.8)	B	18.27	9.87	12.96
	D	15.72	8.11	8.67
III (291.4 and 8.8)	B	15.02	18.71	29.27
	D	15.29	19.68	37.30
IV (90.9 and 2.8)	B	20.00	22.24	20.42
	D	16.68	18.74	13.43
V (291.4 and 2.8)	B	17.91	23.23	27.54
	D	16.07	20.96	23.16

[a] The numerical values given next to each numbered group reflect the intensity values of the conflict stimuli.
[b] The designation (B) and (D) represents the Bright and Dim instructional condition, respectively.

Subjects. The subjects were 70 volunteer females from introductory psychology who were randomly assigned, 14 per group, to one of five groups. Within each group the brighter of the two stimuli was designated as the correct stimulus for half of the group and the dimmer for the remaining half.

RESULTS

The mean speed scores (1/RT × 1000) for all subgroups during the three major phases of the experiment are presented in Table 1. Figure 1 contains the data for the five major conflict groups with the bright-dim conditions combined over successive eight trial blocks. The data for the various groups were analyzed and will be discussed separately for the three major phases of the experiment.

Preconflict performance. In order to determine whether the subjects in the conflict

groups demonstrated comparable performances prior to exposure to conflict, a Lindquist (1953) Type III analysis of variance was performed on the preconflict training performances. The summary of this analysis appears in Table 2 and indicates that there were only two significant effects; that for the bright-dim instructional condition was significant at less than the .05 level, while the trials effect was significant well beyond the .001 level. Of particular interest is the absence of significant performance differences between the groups who were subsequently given the differential conflict treatments. The finding that speed increases with trials has been obtained many times before and indicates that performance becomes more rapid with practice. On the other hand, the finding of a significant difference between the bright-dim instructional conditions is of some interest. An examination of Table 1 indicates that there is a strong tendency for the subjects who were instructed to respond to the brighter stimulus to be faster. This result is in accord with a portion of the results obtained in a previous study (Castaneda & Worell, 1961). It would appear that through previous experience individuals may learn to respond to the "brighter" stimuli when relatively bright stimuli are present, and that forcing them to respond to dimmer stimuli may arouse interfering or incompatible tendencies. In other words, the instruction to respond to the dimmer stimulus seems to have increased the degree of conflict.

Conflict training. The mean speeds of performance for the five conflict groups are presented in the middle portion of Table 1. A second Lindquist (1953) Type III analysis was performed on the data during conflict training and is summarized in the middle portion of Table 2. This analysis provided two significant effects; the main effect of conflict was significant well beyond the .001 level as was the interaction of trials and the conflict conditions. An examination of the performances of the conflict groups during conflict training in Figure 1 indicates quite clearly the greater speeds of the low and intermediate conflict groups when compared with the two high conflict groups. In relation to the significant trials and conflict interaction, a further inspection of Figure 1 shows that while the low and intermediate conflict groups tended to increase in speed over trials, the high conflict groups tended to become slower. The foregoing results support the view that the conflict training conditions produced marked speed differences between the various groups.

A further analysis was made of the simple effects to determine whether the absolute strengths of competing tendencies provided significant differences in addition to those that are apparent for the relative strengths of these tendencies. It will be recalled that the two high conflict groups were exposed to either two equally bright or two equally dim stimuli. Since the difference in strength between the two stimuli was the same for either

TABLE 2

SUMMARIES OF ANALYSES OF VARIANCE

Source	Preconflict performance			Conflict training			Postconflict performance		
	MS	*df*	*F*	*MS*	*df*	*F*	*MS*	*df*	*F*
Instructions (I)	13,353.64	1	4.73*	8,771.38	1	—	21.69	1	—
Conflict (C)	2,306.62	4	—	130,938.53	4	24.38***	7,620.97	4	3.98**
I × C	1,564.40	4	—	1,796.17	4	—	441.82	4	—
Error (b)	2,825.01	60		5,370.64	60		1,914.47	60	
Trials (T)	60,108.72	1	271.29***	38.32	2	—	4,812.09	2	21.91***
T × I	33.71	1	—	30.57	2	—	322.41	2	—
T × C	174.32	4	—	933.33	8	5.69***	46.35	8	—
T × I × C	292.94	4	—	157.96	8	—	431.09	8	—
Error (w)	221.57	60		164.15	120		219.58	120	

* $p = < .05$.
** $p = < .01$.
*** $p = < .001$.

pair of stimuli, the major difference between the conditions is in terms of the absolute intensity of the stimuli. Likewise, one pair of intermediate conflict stimuli was made up of the 291.4 and 8.8 brightnesses, while the other pair consisted of the 90.9 and 2.8 brightnesses. In each case, there was about a 1.5 log unit difference between the stimuli. Assuming that within the present range of values, equal log differences represent subjectively equal differences, the major difference between the two intermediate conflict conditions also is in terms of the absolute intensity of the stimuli. Analyses revealed that the differences between the two high conflict groups and between the two intermediate conflict groups were not significant. Thus, there is no evidence for an effect of the absolute intensities on conflict performance.

Postconflict performance. The principal interest of this study was in the effects of prior exposure to different levels of conflict on performance in weak conflict situations. In order to determine these effects, an analysis was performed on the difference scores for each individual between preconflict performance (Trials 9–16) and postconflict performance (divided into successive eight trial blocks). The mean values for postconflict minus preconflict performance are presented in Table 1. The summary of an analysis of variance performed on these data may be seen in the third column of Table 2. This analysis provides two significant effects; the effect of trials was significant beyond the .001 level and the effects of conflict beyond the .01 level. Once again, the trials effect indicates that performances tend to become more rapid with practice (Figure 1). Of particular interest, however, is the significant conflict effect which suggests that previous exposure to various levels of conflict led to significant differences in the speed of resolving weak conflict situations. Reference to Table 1 and Figure 1 indicates that the two high conflict groups were slower than the two intermediate and lone low conflict groups. Moreover, the differences between these groups tend to persist over the entire 24 trials that (the weak conflict condition) were presented in Figure 1.

DISCUSSION

The principal findings provide clear support for the view that the effects of conflict generalize to other related but less conflictful situations. It was found that individuals whose behavioral efficiency was impaired by exposure to relatively severe conflicts also demonstrated impaired ability to resolve subsequent weaker conflicting tendencies. When the present results are contrasted with those obtained in previous work (Worell & Castaneda, 1961a), there is a strong indication that generalization of the effects of conflict occurs along dimensions of similarity. Thus, in the former study where there was a minimal similarity between the conflict situation and the learning conditions, the generalization of conflict effects was only manifest when one took into consideration the nature of the individual's response to conflict. In the present study where there was a relatively high degree of similarity between the conflict situation and the subsequent performance task, a striking effect of conflict on later performance in a weaker conflict situation was found. From the standpoint of clinical approaches, these findings are in accord with what would be expected if one views personality as being a function of both the individual and the situation. Thus, the tendency to characterize an individual as generally indecisive may be adequate for a few persons, but that better prediction of behavior might be achieved by taking into account the specific nature of previously experienced conflicts and the present conflict situation.

The behavior of subjects in weak conflict following exposure to varying severities of conflict is consistent with expectations derived from the competing response formulation. It will be recalled that this hypothesis proposed that there would be a generalization of competing responses from the conflict training situation to the weak conflict situation, such that the behavior of subjects in weak conflict following performance under strong conflict would be significantly impaired. At the present time the exact nature of these hypothesized competing responses is unknown. Such responses, however, as hesitancy or holding back a response or learning to spend more time in looking from one alternative to the

other appear to be likely possibilities. Although the identification of these responses would be of some value, the consistency of the obtained data with the general competing response position is of most immediate importance.

Although the findings support the competing response formulation, they do not eliminate the importance of the dynamogenic view, or at least a modified version of this view. For example, it might be maintained that the greatest amount of emotionality or drive was aroused in both high conflict groups, but that this heightened emotionality acted in concert with the competing response tendencies. Thus, the tension aroused by the conflict conditions might be seen as having increased the general level of drive and thereby augmented the strengths of the competing tendencies. Such a "feedback" view has been discussed by both Brown and Farber (1951) and Worell (1961). Although admittedly speculative, this reasoning does serve to highlight our lack of independent evidence about the degree of emotionality engendered in any of our groups. In any event, whatever the status of conflict produced drive may be under the present circumstances, it seems apparent that a primary place must be made for some sort of associative (and in this case a competing response) formulation.

Apart from the competing response and dynamogenic views, an alternative explanation based on failure may be considered. It is possible that the various conflict training conditions were seen as differentially difficult by the subjects and that this varying difficulty produced differing levels of feelings of failure. In relation to this point, it should be noted that an attempt was made to minimize the possibility of failure occurring by having every response made by the subject to the conflict situation be a correct one. Thus, at all times the conflict situation was removed by the subject pressing the button under either simultaneously exposed light. Even if feelings of failure were not eliminated by this procedure, however, all that a precisely defined "failure" hypothesis would provide in terms of prediction would be a specification of the kinds of competing responses aroused as a function of conflict. In short, the general predictions made in this experiment would not differ from those made with the competing response formulation, since the "failure" hypothesis would be a special case of the competing response formulation.

In addition to these considerations, the findings relevant to the role of the absolute and relative strengths of conflicting tendencies deserve some attention. Our results suggest that the strength or degree of conflict varies with the *difference* in strength between competing tendencies. Hence, the more closely competing tendencies approximate one another in strength the greater the degree of conflict. On the other hand, we have found that the absolute strengths do not contribute significantly to performance. Here, subjects exposed to either high conflict condition or to either intermediate conflict condition did not differ significantly in speed. As a result, within the present definition of absolute strength, strength of conflict appears to be principally, if not exclusively, based on the difference in strength between competing tendencies. Aside from the fact that other ways of manipulating absolute strength may provide different results, it is worth remembering that the present conflict conditions might be regarded as primarily involving conflict between two approach responses. Whether conflict situations which contain avoidance tendencies will provide similar results remains to be seen.

SUMMARY

As a partial outgrowth of the frequent clinical observation that individuals who are faced with strong conflicts show inefficient handling of these conflicts and weaker conflicts as well, the problem of the generalization of the effects of conflict was investigated. The first question that was examined was whether prior exposure to conflict affected performance in a new but similar conflict situation. In order to anticipate performance in a conflict situation following differential conflict training, two alternative hypotheses, termed the competing response and dynamogenic, were considered. According to the competing response view, exposure to conflict was assumed to lead to the learning of certain conflict-specific responses and performance in a new but similar conflict situation was ex-

pected to be affected by the generalization of these responses from previous conflict experiences. According to the dynamogenic view, conflict was assumed to generate drive and performance in a subsequent conflict situation would be affected as a function of augmented drive levels. The second interest of this study was in the problem of the degree or strength of conflict. An attempt was made to determine the role of the absolute and relative strengths of conflicting tendencies in performance.

To test these considerations, five groups of subjects were exposed to three successive phases of an experimental procedure. Initially, all subjects were exposed to a weak conflict situation. Following this, the subjects in the various groups were given training in responding to one of five levels (severities) of conflict. Finally, all subjects were again exposed to the original weak conflict situation. The findings indicated (*a*) that prior exposure to strong conflict led to a significant impairment of speed performance in weak conflict, and (*b*) that the degree of conflict was dependent on the relative strengths of or the difference in strength between competing tendencies and not on the absolute strengths of these tendencies. The significance of these results for our hypotheses and personality predictions was discussed.

REFERENCES

ANDREAS, B. B. Motor conflict behavior as a function of motivation and amount of training. *J. exp. Psychol.*, 1958, **55**, 173–178.

BERLYNE, D. E. Conflict and choice time. *Brit. J. Psychol.*, 1957, **48**, 106–118.

BERLYNE, D. E. *Conflict, arousal, and curiosity*. New York: McGraw-Hill, 1960.

BILODEAU, INA MCD. Conflict behavior as a function of the strengths of competing response tendencies. Unpublished doctoral dissertation, State University of Iowa, 1950.

BLOCK, J., & PETERSEN, P. Some personality correlates of confidence, caution, and speed in a decision situation. *J. abnorm. soc. Psychol.*, 1955, **51**, 34–41.

BROADHURST, P. L. Emotionality and the Yerkes-Dodson law. *J. exp. Psychol.*, 1957, **54**, 345–352.

BROWN, J. S. Factors determining conflict reactions in difficult discriminations. *J. exp. Psychol.*, 1942, **31**, 272–292.

BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.

CARTWRIGHT, D. Relation of decision-time to the categories of response. *Amer. J. Psychol.*, 1941, **54**, 174–193.

CASTANEDA, A., & WORELL, L. Differential relation of latency and response vigor to stimulus similarity in brightness discrimination. *J. exp. Psychol.*, 1961, **61**, 309–314.

FINGER, F. W. Quantitative studies of "conflict": I. Variations in latency and strength of the rat's response in a discrimination-jumping situation. *J. comp. Psychol.*, 1941, **31**, 97–127.

HOVLAND, C. I., & SEARS, R. R. Experiments on motor conflict. I. Types of conflict and their modes of resolution. *J. exp. Psychol.*, 1938, **23**, 477–503.

HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.

KAUFMAN, E. L., & MILLER, N. E. Effect of number of reinforcements on strength of approach in an approach-avoidance conflict. *J. comp. physiol. Psychol.*, 1949, **42**, 63–74.

KORMAN, M. Ego strength and conflict discrimination: An experimental construct validation of the ego strength scale. *J. consult. Psychol.*, 1960, **24**, 294–298.

LEWIN, K. Environmental forces. In C. Murchison (Ed.), *Handbook of child psychology*. Worcester, Mass.: Clark Univer. Press, 1933. Ch. 14.

LINDQUIST, E. R. *Design and analysis of experiments in psychology and education*. New York: Houghton Mifflin, 1953.

MILLER, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. 1. New York: Ronald, 1944. Ch. 14.

MILLER, N. E., & STEVENSON, S. S. Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery. *J. comp. Psychol.*, 1936, **21**, 205–231.

SCODEL, A., RATOOSH, P., & MINAS, J. S. Some personality correlates of decision making under conditions of risk. *Behav. Sci.*, 1959, **4**, 19–28.

SEARS, R. R., & HOVLAND, C. I. Experiments on motor conflict. II. Determination of mode of resolution by comparative strengths of conflicting responses. *J. exp. Psychol.*, 1941, **28**, 280–286.

WORELL, L. A theory of conflict: Implications for some motivational studies of performances. Unpublished manuscript, Oklahoma State University, 1961.

WORELL, L., & CASTANEDA, A. Individual differences in conflict resolution and the ease of learning. *Amer. Psychologist*, 1961, **16**, 400. (Abstract) (a)

WORELL, L., & CASTANEDA, A. Response to conflict as a function of response-defined anxiety. *J. Pers.*, 1961, **29**, 10–29. (b)

WORELL, L., & HILL, L. Ego strength and anxiety in discrimination conflict resolution. *J. consult. Psychol.*, 1962, in press.

(Received April 10, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 446–449

CRITIQUE AND NOTES

THE EFFECTS OF THREE TYPES OF AGGRESSION ON VASCULAR PROCESSES [1]

JACK E. HOKANSON AND MICHAEL BURGESS

Florida State University

Several recent investigations (Hokanson & Shetler, 1961; Worchel, 1957) have presented findings which suggest that postfrustration expression of aggression reduces general physiological arousal. Worchel, using verbal aggression, inferred this reduction in arousal from elevated postaggression scores on an intellectual task; whereas Hokanson and Shetler measured the course of systolic blood pressure over frustration and physical aggression. The purpose of the present study is to investigate more systematically the variable, type of aggression. To this end, systolic blood pressure and heart rate are measured before and after frustration and after each of four aggression conditions: Physical, Verbal, Fantasy, and No Aggression.

METHOD

Eighty (56 female and 24 male) white, college age subjects were used in a 2×4 factorial design experiment with 3 males and 7 females in each cell. Half of the subjects were frustrated by the experimenter and subsequently placed in one of four aggression conditions: expressing Physical, Verbal, Fantasy, or No Aggression. The remaining 40 subjects were in a nonfrustrated control group and subsequently were also placed in one of the four aggression conditions. Subjects were assigned to conditions on a random basis with the restriction of maintaining equal male-female proportions in each cell. Heart rate was measured continuously and systolic blood pressure at about 2-minute intervals over the entire procedure, with the critical measurements being before and after frustration and immediately after opportunity to aggress.

Subjects were introduced to the experiment as one involving physiological responses to performance on intellectual tasks. To aid in establishing this deception, subjects were administered the Picture Completion subtest of the WAIS following an initial 8-minute adaptation period. During the second "task" of the session, the frustration conditions were introduced. The 40 subjects in the High Frustration condition were asked to count backwards from 100 to 0 by two's as quickly as they could. Following a standardized procedure during this counting, the experimenter repeatedly interrupted and harassed the subject for counting too slowly, having the subject start over three times. After 1 minute the experimenter asked that the counting be stopped, stating that this data could not be used owing to the subject's uncooperative attitude. Subjects in the No Frustration condition were allowed to count backwards to zero without interruption, with the experimenter remarking "Good" at the completion of the series.

Immediately following the frustration procedures, the aggression phase of the experiment took place. The 20 subjects in the Physical Aggression condition were instructed as follows:

> The next task involves an interpersonal guessing problem. You are to think of a number between one and ten and I [the experimenter] will try to guess the number. If I am wrong, you are to depress this plunger which will administer an electric shock to me. If I am correct, simply nod, and go on to the next trial. In this way I will be studying the effects of physical pain on guessing behavior.

The subject was given 10 opportunities to shock [2] the experimenter with subjects in both the High and Low Frustration conditions depressing the plunger an average of 8 times.

The 20 subjects in the No Aggression condition went through exactly the same procedure as described above with one exception: when they depressed the plunger, a small light was turned on which, they were told, signaled to the experimenter his error. Again, each group signaled errors an average of eight times.

The Verbal Aggression condition consisted of having the subject fill out a brief questionnaire evaluating the experimenter's capabilities as an experimenter, in the experimenter's presence. A typical item was to rate the experimenter's competency in relating to subjects.

Finally, 20 subjects were asked to create a story to Card 8BM of the TAT. This card was selected for the Fantasy Aggression condition owing to its moderate picture-pull towards aggressive themes (Eron, 1950).

Each aggression condition took 2 minutes, following which systolic blood pressure and heart rate were again noted. Blood pressure was measured via a sphygmomanometer operated by the experimenter.

[1] This project was supported by the Research Council of Florida State University.

[2] Although the experimenter was wired to an elaborate "shock" apparatus, he did not actually receive a shock, but merely behaved as if he had.

The heart rate apparatus consisted of a rubber diaphragm transmitting pulses from the carotid artery through a pneumatic system, with a pen marking pulses on a constant speed kymograph.

RESULTS

Analyses of the changes in systolic blood pressure and heart rate following the frustration manipulations indicate that the High Frustration condition has significantly higher elevations on both measures. High Frustration subjects manifest a mean increase of 10.3 millimeters of Hg and 9.0 beats per minute, while subjects in the Low Frustration condition increase on the average only 3.2 millimeters and .75 beats per minute, respectively. Both of these differences are significant at the .01 level.

Table 1 presents the analysis of variance and results of the Duncan multiple range test for change in systolic blood pressure after the aggression opportunities, with prefrustration base level as reference.

The significant main effect for Frustration indicates that relative to subjects in the Low Frustration condition, frustrated subjects manifest a significantly more elevated systolic blood pressure immediately after the aggression phase of the experiment (5.8 vs. 2.9 millimeters of Hg above prefrustration base level).

The significant F ratios for Aggression and the Frustration × Aggression interaction may best be interpreted by reference to the cell means and range test results at the bottom of Table 1. The systolic elevations, re base level, of subjects under High Frustration-Physical Aggression (1.2) and High Frustration-Verbal Aggression (1.8) are not significantly different from the elevations in their respective Low Frustration control cells (3.2 and 3.4). On the other hand, the systolic elevations of subjects under the High Frustration-Fantasy Aggression (9.2) and High Frustration-No Aggression (10.8) are significantly greater ($p = .05$) than the elevations in their control conditions (2.2 and 2.8) and in the High Frustration-Physical and Verbal Aggression cells.

Reference to Table 2 and the cell means at the bottom of the table suggests substantially the same pattern of results with the heart rate measure. Frustrated subjects who were given a chance to physically or verbally aggress manifested returns in heart rate to +.8 and −3.8 beats per minute with respect to base level. These values are not significantly different from the rates in the Low Frustration control conditions (−1.2 and +2.8 beats per minute, respectively). Cell means under High Frustration-Fantasy Aggression (8.0) and High Frustration-No Aggression (11.4) are again significantly ($p = .05$) higher than the means in either their respective Low Frustration control conditions (−1.8 and −.2 beats per minute) or the High Frustration-Physical and Verbal Aggression cells (−.8 and −3.8 beats per minute, respectively).

To summarize the vascular changes: frustrated subjects manifest a mean increase of 10.3 milli-

TABLE 1

ELEVATION OF SYSTOLIC BLOOD PRESSURE AFTER AGGRESSION, RE PREFRUSTRATION BASE LEVEL

Source	*df*	*SS*	*MS*	*F*
Frustration	1	163	163.0	10.00*
Aggression	3	311	103.7	6.36*
Frustration × Aggression	3	434	144.7	8.88*
Within	72	1176	16.3	
Total	79	2084		

Cell Means of Elevation in Systolic Pressure

Condition	Aggression			
	Physical	Verbal	Fantasy	None
Low Frustration	3.2_a	3.4_a	2.2_a	2.8_a
High Frustration	1.2_a	1.8_a	9.2_b	10.8_b

Note.—Cells with different subscripts are significantly different from one another at the .05 level, by Duncan multiple range test.
* $p = .01$.

TABLE 2

ELEVATION OF HEART RATE AFTER AGGRESSION, RE PREFRUSTRATION BASE LEVEL

Source	*df*	*SS*	*MS*	*F*
Frustration	1	352.8	352.8	9.28*
Aggression	3	505.0	168.3	4.43*
Frustration × Aggression	3	1038.2	346.1	9.11*
Within	72	2736.0	38.0	
Total	79	4632.0		

Cell Means of Elevation in Heart Rate

Condition	Aggression			
	Physical	Verbal	Fantasy	None
Low Frustration	-1.2_a	2.8_{ab}	-1.8_a	-0.2_a
High Frustration	0.8_a	-3.8_a	8.0_{bc}	11.4_c

Note.—Cells with different subscripts are significantly different from one another at the .05 level, by Duncan multiple range test.
* $p = .01$.

meters of Hg and 9.0 beats per minute as a result of the frustration manipulation; frustrated subjects given an opportunity to physically or verbally aggress against the frustrator show a return of systolic pressure and heart rate to levels not significantly different from their Low Frustration control levels; whereas subjects given an opportunity to aggress in fantasy, or given no opportunity, maintain significantly elevated pressures.

DISCUSSION

The results of this study offer support for the hypothesis that the expression of aggression towards a frustrator reduces frustration produced physiological arousal. Several considerations, however, tend to limit the generality of this hypothesis. Type of aggression apparently is an important variable owing to the failure of the fantasy aggression condition to lower the elevated vascular processes. While the time allowed for fantasy aggression was relatively short (2 minutes) in this study, it may be said that within this period opportunity for the direct communication of physical or verbal aggression to the frustrator is substantially more efficient in reducing systolic pressure and heart rate.

Contrary to attempts by previous authors (Feshbach, 1955; Worchel, 1957), the present investigators caution against concluding that the results reflect a reduction in "aggressive drive." Indeed, the current research was instituted in an effort to overcome the methodological difficulties encountered when attempting to demonstrate such drive diminution. Feshbach's finding that "residual aggression" towards a frustrator (measured by Sentence Completion test and questionnaire) is reduced by interpolated opportunity for fantasy aggression has been questioned by Berkowitz (1958). Owing to the possible arousal of guilt and/or aggression anxiety on the part of his fantasy aggression subjects, Feshbach's results could be explained on the basis of response suppression, rather than drive reduction. Certainly, the failure of the current investigation to find any arousal reduction with fantasy aggression raises further questions about Feshbach's conclusions.

Worchel (1957), on the other hand, assumes a reduction in "aggressive drive" on the basis of heightened intellectual performance with his "opportunity to (verbally) aggress" group. While tenable, this conclusion is at least one step away from his data; and a more parsimonious explanation would be that performance is increased due to general "tension" reduction. This latter conclusion is consistent with the present data involving verbal aggression.

With respect to the current study it should be noted that both the motor activity involved in expressing aggression and the time allowed for expressing aggression can be ruled out as factors affecting the course of the vascular processes. In the Physical Aggression and No Aggression conditions, the physical activity demanded of the subject was the same, the only difference being that in the former (arousal reducing) condition the subject was aware that he was physically hurting the experimenter. Along the same lines, in the verbal and fantasy aggression conditions, the subject engaged in approximately the same amount of writing, with again only the former condition involving the direct communication of aggression to the experimenter. In addition, since the time allowed for each type of aggression was 2 minutes, any effects due to homeostatic reduction of vascular processes would be the same for each condition. It would appear then that the critical factor in the arousal reduction was the subject's knowledge that he was hurting the frustrating experimenter either physically or verbally.

A final point of interest is a comparison of the amount of aggression expressed between frustrated and nonfrustrated subjects for each aggression condition. Under physical aggression, both number of shocks administered and average pressure exerted on the plunger were measured. (Both these measures were found to be valid indicators of increased aggression in a previous study by Hokanson, 1961.) In the Verbal Aggression condition, the degree of negative evaluation of the experimenter was measured by a three-item Likert-type questionnaire. Fantasy aggression was measured via a TAT hostility scale previously successfully used by Hokanson and Gordon (1958). For each type of aggression, the difference in amount or intensity of aggression expressed between frustrated and nonfrustrated subjects was not significant. This rather unexpected finding raises a question of whether the degree of aggression expressed is related to the "cathartic" effects found. A similar question was posed recently by Rosenbaum and deCharms (1960) in a study which found that frustrated subjects expressed only a meager degree of aggression towards the frustrator but still exhibited reduced residual aggression. While the current data do not allow any unequivocal conclusions on this point, perhaps the critical feature in this paradigm is merely *having the opportunity* to aggress towards the frustrator following a frustration.

SUMMARY

Eighty college age subjects underwent either a frustrating or nonfrustrating experience and subsequently were placed in one of four aggression conditions: Physical, Verbal, Fantasy, or No Aggression. The intent of this 2×4 factorial design experiment was to measure the effects of each type of aggression on frustration produced elevations in systolic blood pressure and heart rate.

The results indicated that frustrated subjects who were given an opportunity to physically or verbally aggress against the frustrator, manifested returns on both physiological measures to the levels of the Low Frustration control subjects. Frustrated subjects in the Fantasy and No Aggression conditions exhibited significantly elevated systolic pressures and heart rates at the postaggression recording of these measures.

The results were discussed as offering support for the hypothesis that the expression of Physical or Verbal aggression towards a frustrator tends to reduce general, physical arousal.

REFERENCES

BERKOWITZ, L. The expression and reduction of hostility. *Psychol. Bull.*, 1958, **55**, 257–283.

ERON, L. D. A normative study of the thematic apperception test. *Psychol. Monogr.*, 1950, **64**(9, Whole No. 315).

FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, **50**, 3–11.

HOKANSON, J. E. The effects of frustration and anxiety on overt aggression. *J. abnorm. soc. Psychol.*, 1961, **62**, 346–351.

HOKANSON, J. E., & GORDON, J. E. The expression and inhibtion of hostility in imaginative and overt behavior. *J. abnorm. soc. Psychol.*, 1958, **57**, 327–333.

HOKANSON, J. E., & SHETLER, S. The effect of overt aggression on physiological arousal. *J. abnorm. soc. Psychol.*, 1961, **63**, 446–448.

ROSENBAUM, M. E., & DECHARMS, R. Direct and vicarious reduction of hostility. *J. abnorm. soc. Psychol.*, 1960, **60**, 105–111.

WORCHEL, P. Catharsis and the relief of hostility. *J. abnorm. soc. Psychol.*, 1957, **55**, 238–243.

(Received March 20, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 450–455

THE EFFECT OF AFFECTIVE TONE ON THE VERBAL BEHAVIOR OF PROCESS AND REACTIVE SCHIZOPHRENICS [1]

ALAN S. DEWOLFE [2]

Northwestern University

Increasing attention is currently being paid to the classification of schizophrenics into the so-called Process and Reactive groups. The former term usually denotes a poor premorbid adjustment and the latter a relatively adequate premorbid adjustment.[3] Unlike groups classified by the Kraepelinian system, Reactive and Process groups have been found to differ both prognostically, Reactives having the more favorable prognosis (Kantor, Wallner, & Winder, 1953; Phillips, 1953; Rodnick & Garmezy, 1957; and others), and behaviorally (Becker, 1956; Bleke, 1955; Harris, 1957; among others).

Purely behavioral studies have yielded superficially contradictory results with the Reactives sometimes more efficient, accurate, etc., than the Process group and sometimes vice versa. Physiological studies, however, have yielded more consistent results. Using physiological measures, King (1958) using mecholyl, a sympathetic nervous system depressant, and DeVault (1955) using situational stress, emotional pictures, found the Reactives showed a more intense reaction than the Process group who responded minimally, if at all. More recently, Zuckerman and Grosz (1959) have reported results contradictory to King, but studies of schizophrenics classified as Remitter or Nonremitter, a distinction expected to be correlated to a relatively high degree with the Process-Reactive classification, by Funkenstein, Greenblatt, and Solomon (1952) and Hirschstein (1955) found results similar to King's and thus provide indirect support for his findings. These physiological studies suggested the generalization that the Reactive schizophrenics' response to affective stimulation is greater than that of the Process' response and is usually equal to or greater than the Normals' response.

The physiological findings further suggest that the previously mentioned inconsistencies in behavioral level comparisons of Process and Reactive groups may be the result of differences in reactivity. Attempts by the present writer to classify tasks and/or situations as stressful or affective led to the hypothesis that Reactives tend to be superior in neutral situations (e.g., Becker, 1956; Bleke, 1955; Harris, 1957) but inferior in affective situations. Studies by Williams and Machi (1957) and Zubin, Windle, and Hamwi (1953) using the Remitter vs. Nonremitter classification found results to support both the Reactives' superior performance on neutral tasks, perceptual-motor tasks and rote learning of nonmeaningful material, and inferior performance in situations with affective connotations (poorer performance on intellectual and memory tasks and slower reaction times in word association tests).

The less efficient performance of Reactives in stress or affective situations may be due to interference phenomena, emotional responsiveness leading in some way to response tendencies which interfere with performance, as is sometimes the case with normal subjects. The Reactives being more responsive, like Normals, may, therefore, exhibit more interference effects than the less responsive Process group. Tasks unfortunately are not always clearcut on affectivity-nonaffectivity and their results, therefore, provide only suggestive evidence for the hypothesis outlined above.

The present experiment was designed to investigate with a purely behavioral task, the hypothesis that Reactive schizophrenics are more responsive to affective stimuli than Process schizophrenics and that this emotional respon-

[1] This article is based on a doctoral dissertation submitted to the Graduate School of Northwestern University. The author wishes to express his sincere appreciation to Janet A. Taylor for her aid, advice, and encouragement during the investigation. He is also indebted to the departments of psychology and psychiatry at the Downey Veterans Administration Hospital, the Hines Veterans Administration Hospital, and the Chicago West Side Veterans Administration Hospital, for their cooperation in making subjects and facilities available.

[2] Now at the Psychology Service, Veterans Administration Hospital, Hines, Illinois.

[3] Investigators have not uniformly agreed as to whether there are two classes of schizophrenia or whether the groups are polarities of a continuum. Lack of clearcut bimodal distribution in any of the measuring systems makes either view possible. However, regardless of the investigator's predilections, research has typically involved selection and investigation of the two extreme groups. Usage of the terms Process and Reactive in the present study is meant only to reflect this procedure and not necessarily advocation of the position of separate types of schizophrenia.

siveness may in certain types of tasks lead to response interference. The behavioral response measures consisted of the personal pronoun "I" as a measure of personal reference and reaction time in making up sentences. The personal reference measure was intended to show the degree of emotional responsiveness as judged by an avoidance reaction, i.e., a decrease in personal reference under the threat of the affectively toned verbs, while the reaction time measure was invoked to demonstrate that emotional reactivity may lead to response interference and "poorer" performance, i.e., increased reaction time. According to speculations offered above, the greater reactivity of the Reactives was expected to result in the greatest *decrease* in rate of the use of "I" (avoidance response) and the greatest *increase* in reaction time (interference effect) with the affective verbs in the Reactive group. The smallest differences were anticipated in the Process group and an intermediate magnitude of difference in the Normal group on both measures. The order of the effect of affective stimuli was expected to be the same in both measures with the direction of change opposite in the two measures, i.e., a decrease in personal reference and an increase in reaction time. The greatest effect was anticipated in the Reactives, the least in the Process group, with the Normals showing a magnitude of effect between two schizophrenic groups on both measures.

METHOD

Subjects.[4] Two schizophrenic groups and one Normal Control group, each consisting of 15 male inpatients from Veterans Administration hospitals were used. All subjects, schizophrenic or normal, were between the ages of 20 and 55 with a minimum of an eighth grade education, a current hospitalization of under 9 months duration, and by history had neither inherited nor acquired organic brain pathology.

The schizophrenic subjects met these additional criteria: no more than two nonmilitary psychiatric hospitalizations including the current hospitalization and no form of convulsive or coma therapy for 9 months prior to the experiment. Schizophrenic subjects all had an established diagnosis of schizophrenic reaction and were classified as Process or Reactive by scores on the Phillips (1953) Scale of Premorbid Adjustment in Schizophrenia. The Process group scored 18 or over on the scale and the Reactives 14 or under. The Normal Control subjects were selected from the orthopedic and tuberculosis services and had not been hospitalized more than 9 months in the 2 years prior to the current hospitalization and had no history past or present of significant abnormal emotional response.

In the three groups, the subjects were individually matched for age, educational level, and socioeconomic level. Educational level was scaled on three intervals: not a high school graduate, high school graduate, high school graduate with more than one semester of college. Socioeconomic status was scaled by occupational level in a three-level modification of the Centers (1949) index.[5] The subjects from each major group were assigned in counterbalanced order to one of two subgroups which differed only in which neutral task was presented first. All subjects who successfully made up the 150 sentences were used. Any sentence using one of the pronouns presented was accepted regardless of the accuracy of perception of the stimulus verb or the grammatical correctness or meaningfulness of the sentence.

Materials and Apparatus. Stimuli for the experimental task were 150 slides. Each slide contained a different past tense verb in the center and four pronouns (HE, SHE, THEY, and I) on a single line below the verb. All words were in upper-case letters. The slides were projected on a screen by a 35-millimeter projector. The 150 slides were divided into three groups of 50, two groups with verbs neutral in meaning and one with affective verbs. The affective and neutral verbs were selected by six judges—advanced graduate students in psychology or experienced clinical psychologists.

The order of pronouns on the slides was varied systematically and the Affective and Neutral slide groups were equated for order of pronouns. The order of presentation of the slides within the three groups, affective and two neutral, was randomized by shuffling slides within the groups between-subjects. To measure reaction times, the projector was equipped with a single switch which simultaneously activated a clock and projected the slide on the screen. The experimenter stopped the clock with a separate hand switch when the subject responded.

Procedure. Each subject was tested individually in a single period of about an hour. When the subject entered the room and was seated, the experimenter moved to the projector behind the subject and darkened the room sufficiently to make the experimental stimuli readily visible, but did not totally darken the room. Subjects were instructed that they would see slides with a centered verb and four pronouns, that their task was to make up sentences, and that any kind of a sentence was acceptable as long as they used one of the pronouns and the verb. These instructions were projected on the screen. Then, any

[4] Details on subjects and their selection have been deposited with the American Documentation Institute. Order Document No. 7090 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

[5] A more detailed description of the matching and other experimental procedures, as well as more inclusive theoretical discussions can be found in the original dissertation available from University Microfilms, Inc., Ann Arbor, Michigan (L. C. Card No. Mic 6-06542).

TABLE 1

ANALYSES OF VARIANCE OF "I" FREQUENCY DATA

Source	A. All conditions analysis			B. Neutral 1-Neutral 2 analysis		
	df	*MS*	*F*	*df*	*MS*	*F*
Groups	2	362.8	1.27	2	324.1	1.52
Between-subjects same group	42	285.2		42	213.7	
Total between-subjects	44			44		
Conditions	2	207.6	6.68*	1	9.4	0.47
Groups × Conditions	4	42.9	1.38	2	61.8	3.12**
Pooled Subjects × Conditions	84	31.1		42	19.8	
Total within-subjects	90			45		

* $p < .01$.
** $.05 < p < .10$.

misunderstanding about the instructions and any task relevant questions were answered.

The first slide was presented and if any aid was required, the experimenter repeated the instructions. Each subject was given the 150 slides: 50 neutral, then 50 affective, and finally 50 more neutral with the particular set of neutral slides given first counterbalanced within each group.[6] At regular intervals in each set of 50 slides, subjects were encouraged by a general statement like, "That's fine, you're doing well." After every sentence, the subjects were reinforced by the word "ummhmm" spoken by the experimenter regardless of the pronoun used, i.e., the task was used to study the effect of affective tone on performance rather than the conditioning process. The subjects were prepared for the next stimulus by the experimenter saying, "All right, ready," prior to each new slide.

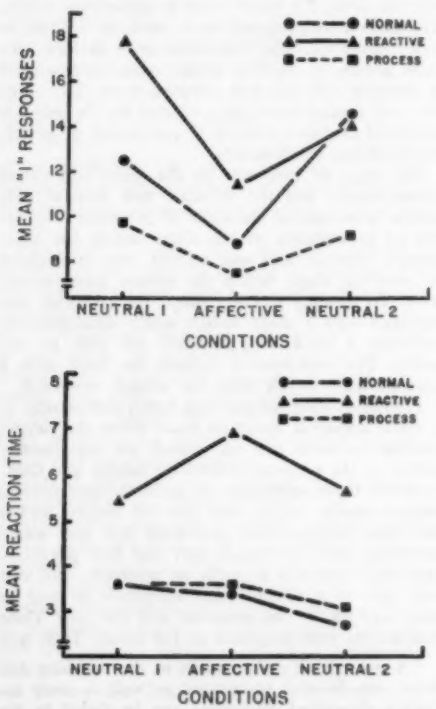

FIG. 1. Number of personal references per 50 trials and mean reaction time in seconds per 50 trials in the Affective and Neutral conditions.

RESULTS

Personal Reference. The personal reference data were evaluated by an analysis of variance with Groups (Reactive, Process, and Normal) and Conditions (Neutral 1, Affective, and Neutral 2) as the main effects. Since all subjects received all three conditions, this variable was treated as the correlated term in an analysis for repeated measures in all analyses in this study. The F of 6.68 for Conditions was significant ($p < .01$) with all others not significant (Table 1). An analysis of variance on the initial neutral condition showed that the groups were far from

[6] Subjects were presented with a group of neutral verbs first and affective second, and never the reverse. This procedure was followed since findings of Rodnick and Garmezy (1957) suggest that if the affectively toned verbs were presented first, the effects of this material might persist during the Neutral condition and depress measures of the difference between the two conditions. Further, the groups might be differentially affected with the Reactives expected to be most responsive and slowest to dissipate their emotional response showing the greatest effect. In order to test for the presence of perseveration affects, as well as to provide a check on the possibility that differences in the Affective condition were due primarily to practice effects, a second group of neutral verbs was added following the affective slides.

significantly different in rate of personal reference in this condition ($F < 1.00$).

In the analysis of the two neutral conditions (Table 1), the F for Groups was not significant, but the Groups × Conditions interaction neared significance ($.05 < p < .10$) which was probably the result of the Normals showing an increase in personal reference between the two neutral conditions (upper portion of Figure 1). In this analysis, the Conditions F was far from significant indicating that differences in the analysis with the Affective condition present were not the result of time or trials. In the analyses of the "I" data, Bartlett tests showed heterogeneity of variance significant between the .01 and .05 levels which was not considered sufficiently extreme to warrant rejection of the conclusions from these analyses.

These analyses verified the hypothesis of the general effect of the affective tone variable. Although the Groups × Conditions interaction F was not significant, some indirect support for the hypothesis of the differential effect of the affective stimuli was found in the fact that the groups differed in the expected direction and fell in the predicted order of magnitude of difference between conditions. Analysis by direct difference t tests within the groups showed that the Reactives ($t = 2.74$) and Normals ($t = 3.08$) had highly significant ($p < .01$) differences between the neutral and affective conditions, while the Process group ($t = 1.31$) did not differ significantly.

Reaction Time. As indicated in the lower portion of Figure 1, the Reactives displayed a sharp increase in mean reaction time in the Affective condition and their mean reaction time in the final Neutral condition was slightly higher than in the initial Neutral condition. The Process and Normal groups showed negligible change in the Affective condition, followed by a slight drop in mean reaction time in the final Neutral condition.

In the initial analysis of variance with reaction time data (Table 2), both main effects and the Groups × Conditions interaction effects were significant. In t test analyses, the group did not differ significantly in the initial Neutral condition. In the Affective and final Neutral conditions, the Reactives had a reaction time significantly higher than the other groups which did not differ from each other. Comparing the two neutral conditions (Table 2), the Groups F was significant ($p < .05$). The Conditions and Groups × Conditions interaction neared significance ($.05 < p < .10$) apparently due primarily to the magnitude of the drop in mean reaction time between the first and final Neutral conditions in the Process and Normal groups (lower portion of Figure 1).

Bartlett tests on the reaction time data uniformly showed extreme heterogeneity of variance (all Bs having a $p < .001$). Therefore, the reaction time data was reanalyzed by distribution free analyses. The Groups variable was evaluated by a Kruskal-Wallis test and Conditions by a Cochran Q test. Neither was significant. As a distribution free analysis equivalent of the Groups × Conditions interaction effect, the groups were compared by a Kruskal-Wallis test using the difference between the first neutral and affective conditions as the raw scores. The H of 7.10 was significant ($p < .05$). The two kinds of analyses then, gave parallel significant results on the predicted differences in reaction time among the groups.

Additional Analyses. The relationship between the two measures—reaction time and frequency of using "I"—was studied by means of Spearman rho's. Using the totals for both measures, the rho for all subjects was $-.36$ ($p < .05$). In analyzing this relationship in the groups taken

TABLE 2

ANALYSES OF VARIANCE OF MEAN REACTION TIME DATA

Source	A. All conditions analysis			B. Neutral 1–Neutral 2 analysis		
	df	*MS*	*F*	*df*	*MS*	*F*
Groups	2	115.0	4.39**	2	58.6	3.78**
Between-subjects same group	42	26.2		42	15.5	
Total between-subjects	44			44		
Conditions	2	7.6	6.90*	1	4.5	3.46***
Groups × Conditions	4	3.5	3.18**	2	3.2	2.46
Pooled Subjects × Conditions	84	1.1		42	1.3	
Total within-subjects	90			45		

* $p < .01$.
** $p < .05$.
*** $.05 < p < .10$.

separately, the rho's were as follows: Reactive $-.60$ ($p < .05$), Normal $-.65$ ($p < .05$), and Process .13 (*ns*).

DISCUSSION

The hypothesis that Reactive schizophrenics are more emotionally responsive than Process schizophrenics was confirmed. The evidence further suggests that the Process schizophrenics are less responsive than Normals. These results using affective verbs with behavioral measures parallel the results of DeVault (1955) who used emotional pictures and a physiological measure, indicating that the greater emotional responsiveness of Reactive schizophrenics appears at both levels—the behavioral and physiological.

The Reactives' significant increase in reaction time in the Affective condition coupled with the Process group's slight decrease supported the speculation that the inferior performance of the Reactives on some laboratory tasks may be due to some interference phenomenon. That the interference phenomenon is associated with differences in emotional responsiveness is supported within this study by the not statistically significant but parallel results (same order of responsiveness) in the personal reference measure and by the significant negative correlation between the two measures. The direction of changes was predicted to be opposite in these measures. The association between interference effects and responsiveness is further supported by the parallel results between the reaction time measure and DeVault's (1955) physiological measure of emotional responsiveness. As indicated earlier, the interference phenomenon hypothesis would account for differences found between Process and Reactive groups by Becker (1956), Bleke (1955), Harris (1957) and others and the lack of these differences under different conditions by Rodnick and Garmezy (1957), as well as explaining the seemingly contradictory results within the studies of Williams and Machi (1957) and Zubin, Windle, and Hamwi (1953) with Remitters and Nonremitters.

It was noted that only the Reactives showed the interference effect (increase in reaction time) with the affective verbs (lower portion of Figure 1). The fact that the Reactives were also higher in reaction time (not significantly, however) in the initial neutral condition potentially raises two conflicting explanations: the hypothesis that change is proportional to the original level of reaction time, or the hypothesis that a certain original level of responsiveness is necessary for interference to occur, or threshold hypothesis.

Although the results of this study were not directly applicable to this problem and, therefore, are inconclusive, there was limited support for the latter hypothesis. When the reaction time in the original Neutral condition was correlated with the *change* in reaction time between initial Neutral and Affective conditions within the three groups, a significant relationship appeared only in the Reactive group. The Reactive group rho of .55 was significant at the .05 level, while the Process group rho of .05 and the Normal rho of $-.35$ were not significant. This supports the threshold hypothesis which, assuming the relationship between responsiveness and reaction time, would predict this relationship in the Reactives, while the proportional model would predict this relationship in all groups. Combining the above and the suggestive personal reference data, the present writer tentatively suggests the possibility of relationships between interference phenomena, emotional responsiveness, and what might be called "personalization" of the experimental situation, "ego involvement," or some similar construct. These speculations, however, must await substantiation from further research.

SUMMARY

The present investigation was designed to study emotional responsiveness in Process and Reactive schizophrenics. Process and Reactive groups, classified by level of premorbid adjustment on the Phillips scale, were compared to each other and a Normal Control group on a task eliciting verbal behavior. Each group consisted of 15 hospitalized male subjects. The subjects were required to make up sentences using a given verb and any one of four pronouns. They were presented with a series of 50 verbs neutral in meaning, followed by 50 affective verbs, and finally, a second series of 50 neutral verbs.

The comparisons of major interest involved differences in reaction time and personal reference (use of "I") measures between neutral and affective conditions. It was hypothesized that the order of groups, when compared on the magnitude of differences between these conditions, from greatest to smallest would be Reactive, Normal, and Process and that these differences reflect differences in responsiveness to affective stimuli in both measures and response interference in the reaction time measure. The same order of magnitude of change was predicted for both the measures, although an increase in reaction times and a decrease in personal reference was expected in the affective condition.

In the main, the hypothesis of differential re-

sponsiveness to affective stimuli was verified. In the personal reference measure, the differences between conditions were in the expected direction with the magnitude of differences in the expected order.

Confirmation of the hypotheses of differential emotional responsiveness and interference phenomena was found in the reaction time analyses. As expected, the Reactive group exhibited a significant increase in reaction time in the Affective condition which perseverated into the final Neutral condition, while the changes in the Normal and Process groups were not significant.

REFERENCES

BECKER, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *J. abnorm. soc. Psychol.*, 1956, **53**, 229–236.

BLEKE, R. C. Reward and punishment as determiners of reminiscence effects in schizophrenic and normal subjects. *J. Pers.*, 1955, **23**, 479–498.

CENTERS, R. *The psychology of social class.* Princeton: Princeton Univer. Press, 1949.

DEVAULT, S. H. Physiological responsiveness in reactive and process schizophrenia. Unpublished doctoral dissertation, Michigan State University, 1955.

FUNKENSTEIN, D. H., GREENBLATT, M., & SOLOMON, H. C. An autonomic nervous system test of prognostic significance in relation to electroshock treatment. *Psychosom. Med.*, 1952, **14**, 347–362.

HARRIS, J. G., JR. Size estimation of pictures as a function of thematic content for schizophrenic and normal subjects. *J. Pers.*, 1957, **25**, 651–671.

HIRSCHSTEIN, R. The significance of characteristic autonomic nervous system responses in the adjustment, change and outcome in schizophrenia. *J. nerv. ment. Dis.*, 1955, **122**, 254–262.

KANTOR, R. E., WALLNER, J. M., & WINDER, C. L. Process and reactive schizophrenia. *J. consult. Psychol.*, 1953, **17**, 157–162.

KING, G. F. Differential autonomic responsiveness in the process-reactive classification of schizophrenia. *J. abnorm. soc. Psychol.*, 1958, **56**, 160–164.

PHILLIPS, L. Case history data and prognosis in schizophrenia. *J. nerv. ment. Dis.*, 1953, **117**, 515–525.

RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957.* Lincoln: Univer. Nebraska Press, 1957. Pp. 109–184.

WILLIAMS, R. J., & MACHI, V. S. An analysis of interpersonal correlations among thirty psychotics. *J. abnorm. soc. Psychol.*, 1957, **55**, 50–57.

ZUBIN, J., WINDLE, C., & HAMWI, VIOLET. Retrospective evaluation of psychological tests as prognostic instruments in mental disorders. *J. Pers.*, 1953, **21**, 342–355.

ZUCKERMAN, M., & GROSZ, H. J. Contradictory results using the Mecholyl Test to differentiate process and reactive schizophrenia. *J. abnorm. soc. Psychol.*, 1959, **59**, 145–146.

(Received April 24, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 456–458

NUMBER OF VERBAL INTERRELATIONSHIPS AS A DETERMINANT OF GROUP SIZE

GEORGE F. CASTORE

Mental Health Institute, Mt. Pleasant, Iowa

Group size has been determined by common sense estimates, preferences of the group psychologists, or by traditional attitudes. This brief note is to present one consideration, number of verbal interrelationships as related to group size. The degree to which group size sets the limit on the possible interrelationships between members during a single session is likely to be related to the amount of individual participation which may be achieved.

Attention is given to the direction of remarks to others rather than the source of a response alone. The value for an individual is assumed to reside in an increase in communication potentials by increasing the range of persons to whom a given person can relate. An individual who extends his relationships to a wider range of persons is compelled to do more than just talk about himself or broadcast in a general way. To continue an established verbal relationship, he is involved in the complexity of understanding what another says, seeing how he is understood by another, dealing with specific individual differences in meanings and interpretations, observing how he influences others and then noting effects they have on him. To increase the range of persons with whom one person communicates is likely to increase the skills of communication such as clarifying, anticipating distortion, etc., and provide a broader cultural context within which he may improve and meet his social needs. The advantage to the group as a whole of increasing the amount of interpersonal participation is that a more varied and, hence, richer supply of verbal interaction arises with more opportunity for each person to find a relationship which meets his particular need of the moment. The greater variation of interpersonal interchange is likely to have the additional educational effects of broadening each person's interests and range of social sensitivity. To the psychologist the advantage lies in possession of a simple and reliable tool of observation.

The aim of a therapeutic group is the improvement of the patient. Conditions which increase rate, depth, or reliability of growth that is deemed desirable would generally be accepted as of value. One of these conditions would be opportunity for verbal expression. Opportunity is related to many facets of group interaction. For example, opportunity has a different significance when it is easy to obtain and when it is difficult. The same condition that makes opportunity easy for all is likely to provide a situation in which two members may hold the floor to the exclusion of others. The emergence of subgroups makes verbal expression more difficult for the individual members of a group. Groups which are too small may find that practices such as establishing of norms, forming subgroups, shunting responsibility to the group leader, and holding to stereotyped patterns of interaction, act along with other behaviors to reduce over-all interaction between members. Large groups possess as many obvious objections.

METHOD

To determine the size of a small interaction group that would have optimal growth value for its members, a measure of the extent to which the members initiate verbal relationships with each other in varied size groups was used. This measure was not the frequency of participation. Two persons may produce high frequency of response. The measure was more one of diffusion of interaction, the number of other members to which each person directed at least one remark during an hour's session. In a group composed of eight members, perfect diffusion would be obtained if each member directed at least one remark to seven other members. For any size group, the number of possible interchanges would be represented by $n(n-1)$. A percentage index measure for each group was simply obtained by the formula

$$\frac{f_o}{n(n-1)}, \text{ and } \frac{\Sigma f_o}{\Sigma n(n-1)}$$

for each mean interval value when combining odd-even size groups.

A simple procedure was used to make the observations. A list was made of the members, sitting in a circle, beginning at the psychologist's left. Each person is given a successive number. When person Number 3 speaks to person Number 5, a 5 is written after the name of Person 3. No more record need be made of Person 3 responding to Person 5 during this meeting. Any other person to whom Number

3 speaks is indicated by writing in the appropriate number when the response occurs. The same procedure is used for recording the responses of each person present, excluding the psychologist. At the end of an hour's session, the record will show the extent of verbal relationships each person has made with the other members in the group.

To be recorded as a directed response, the response should possess one of the following four qualities: mention of the recipient's name; mention of a second person singular "you" in such a way that there is no doubt about the referent; a complete sentence answer to another person's question which was directed to the subsequent speaker; and looking directly at the person with whom he is speaking. The implication here is that only one person among those present can be identified as capable of meeting the needs of the speaker at that moment.

There are also some qualities which definitely exclude remarks from being recorded as directed ones. These are simple Yes–No answers or answers which are incomplete or incoherent sentences. A response which elaborates on the response of another but is not directed to anyone particularly as well as a broadcasting response which is said to the group as a whole or for anyone who chooses to be concerned, likewise is not credited. Where there is the slightest doubt as to whether or not a response is directed to a specific person, the response should not be tabulated. The type of response that may be described as "thinking aloud" most often is undirected.

Groups observed were inpatient therapeutic groups in mental hospital settings. The groups were well-practiced, the members of each group having met together for over a month. Consequently, the results given in this study represent a favorable condition of participation. It would be expected that members of groups in earlier stages of development would be less inclined to direct remarks to so many different persons present.

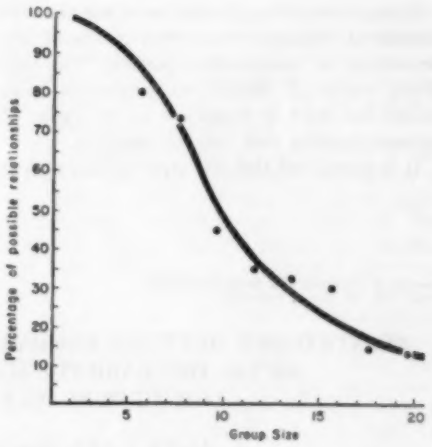

FIG. 1. Mean percentage values representing the extent of relationship initiated during verbal interchange between members of varied sizes.

The age range of the members varied from 14 to 70, the average representing the usual membership of psychotherapeutic groups in state institutions for treatment of severe personality disorders. No records were made on distribution of sex but, since heterogeneity was a feature of group composition, it is roughly estimated that males and females were equally distributed. All types of diagnostic categories were represented in these groups. There was little preponderance of any one type. Although most patients were considered as potentially active treatment patients, a few in each group were regressed. In all, 55 groups were observed with most of the observation falling between groups composed of 11 and 18 members.

TABLE 1

COMPARING THE PERCENTAGE OF MEMBER INTERRELATIONSHIPS AND AVERAGE NUMBER OF INTERRELATIONSHIPS FOR 55 OBSERVATIONS OF VARIED SIZE GROUPS

Group size	Number of groups observed	Average number interrelationships per group	Percentage of relationships possible $\frac{\Sigma f_0}{\Sigma n(n-1)}$
5–6	3	20.0	.800
7–8	11	37.1	.732
9–10	3	33.6	.440
11–12	8	57.8	.348
13–14	11	53.4	.318
15–16	8	57.8	.248
17–18	7	40.5	.141
19–20	4	50.0	.134

RESULTS

There is a marked reduction in the percentage measure, but not the observed number of interactions when group size reaches 9 members. This is shown in Table 1 by comparing average number of interrelationships per group with the percentage for each group size represented. There is also a marked change with 17 or more members present. Based on these data a desirable level of verbal interaction would be determined by the aims of a particular type group. Obviously, for those who are quite reticent and interact with extreme difficulty, more opportunity should arise in a group with less than 9 members. Others could profit from groups having up to 16 members.

Figure 1 shows the distribution of the observed number of initiated relationships translated into percentage of relationships possible. The suggested curve of distribution emphasizes more clearly the effect of group size on the extent of interrelationships that may be expected.

It is recognized that the types of interactions may be a much more important consideration than their frequency of occurrence. Where the increase in participation between members is a desirable feature of group process, the findings presented are likely to be of value in group organization.

(Received April 10, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 64, No. 6, 458–460

RELATIONSHIP BETWEEN PROBABILITY OF ITEM ENDORSEMENT AND SOCIAL DESIRABILITY SCALE VALUE FOR HIGH AND LOW GROUPS ON EDWARDS' *SD* SCALE [1]

ALLEN L. EDWARDS AND JERALD N. WALKER

University of Washington

There is considerable evidence to show that the probability of a True response to a personality item can be predicted quite accurately from a knowledge of the social desirability scale value of the item (Cowen & Tongas, 1959; Edwards, 1953, 1957a, 1957b, 1959; Hanley, 1956). If the items in an inventory are scaled for social desirability, then a socially desirable or socially undesirable response to an item can be defined in terms of the scale value of the item. A socially desirable response is defined as a True response to an item with a socially desirable scale value or a False response to an item with a socially undesirable scale value. A socially undesirable response is defined as a True response to an item with a socially undesirable scale value or a False response to an item with a socially desirable scale value (Edwards, 1957b).

Edwards (1957b) has developed the Social Desirability (*SD*) scale to measure the tendency of subjects to give socially desirable responses to personality items. All of the 39 items in the *SD* scale are keyed for socially desirable responses, but 30 items are also keyed False and only 9 are keyed True. Because of the imbalance in True-False keying of the *SD* scale, Wiggins (1959), Jackson and Messick (1958), and Couch and Keniston (1961) have suggested that the *SD* scale may be measuring an acquiescent tendency rather than a tendency to give socially desirable responses. Evidence has been presented by Edwards (1957b) and Edwards and Walker (1961) which indicates that acquiescence is not an important component of the *SD* scale.

The present study was designed to provide further evidence on the acquiescence and social desirability interpretations of scores on the *SD* scale. If low scores on the *SD* scale reflect the tendency to respond True and high scores the tendency to respond False, then such tendencies should operate over the entire range of the social desirability continuum. Thus, for High and Low *SD* groups, the regression lines of probability of a True response (Y) on the social desirability scale values (X) of a given set of items should, according to the acquiescence interpretation of the *SD* scale, be parallel with the regression line for the Low *SD* group having a higher Y intercept than that for the High *SD* group. Thus, the Low *SD* group should also have a higher mean probability of a True response than the High *SD* group.

A social desirability interpretation of the *SD* scale would result in the same prediction as the acquiescence interpretation with respect to the difference in the Y intercepts. However, the social desirability interpretation of the *SD* scale would predict that the regression lines would not be parallel but rather that they would cross, with the line for the High *SD* group having the steeper slope. The basis for this prediction is that the High *SD* subjects should have a greater probability of a True response than Low *SD* subjects for items with socially desirable scale values. For items with socially undesirable scale values,

[1] This research was supported in part by Research Grant M-4075 from the National Institute of Mental Health, United States Public Health Service.

TABLE 1

MEANS AND RANGES OF SCORES ON THE *SD* SCALE FOR THREE GROUPS OF SUBJECTS

Group	Male *N*	Female *N*	Total *N*	$\bar{X}$	Range
High	42	43	85	36.588	35–39
Middle	22	48	70	32.686	31–34
Low	36	46	82	25.951	12–30

the difference in the probabilities should be reversed with the Low *SD* group having the greater probability of a True response. Furthermore, if scores on the *SD* scale are relatively uninfluenced by the tendency to respond True, then the mean probability of a True response for the High and Low *SD* groups should be much the same.

METHOD

The *SD* scale and 150 items from the EPPS were administered under standard instructions to 237 students at the University of Washington. Subjects were divided into three groups on the basis of their *SD* scores. Records from both sexes were pooled, since there is evidence (Edwards, 1957b) to show that little distortion is to be expected with such pooling. Table 1 gives the composition of each of the groups with their respective means and ranges on the *SD* scale.

Social desirability scale values of the 150 EPPS items were available from a previous study (Edwards, 1953). For each *SD* group, the probability of endorsement for each of the 150 EPPS items was obtained and these probabilities were correlated with the social desirability scale values of the items.

RESULTS AND DISCUSSION

Table 2 gives the means and variances of the probability of endorsement for the 150 EPPS items, for each of the separate groups. The means are somewhat higher than those obtained in another study (Edwards, 1959) in which the items of the Interpersonal Check List (ICL) were used.[2] The difference between the means in the two studies can be attributed to the differences in the social desirability scale values of the two sets of items. The set of EPPS items has a greater proportion of items at the desirable end of the social desirability continuum, than does the set of ICL items. Thus, we would expect to obtain a higher mean probability of endorsement for the EPPS items than for the items of the ICL.

According to the acquiescence interpretation of the *SD* scale, the Low *SD* group should have a higher mean probability of endorsement of the EPPS items than the High *SD* group, whereas the actual difference is in the reverse direction. In accordance with the social desirability interpretation of the *SD* scale, the difference between the means for the Low and High *SD* groups is not statistically significant ($p > .70$).

The correlations between probability of endorsement and social desirability scale values of the items are .844, .710, and .731 for the High, Middle, and Low *SD* groups, respectively. The difference between the correlations for the High and Low *SD* groups is statistically significant ($p < .01$).

It is clear, from the values of the regression coefficients for the separate groups, that the regression lines are not parallel but rather that they cross in accordance with the social desirability interpretation of the *SD* scale. As expected, the slope of the line for the High *SD* group is steeper than that for the Low *SD* group. The difference in the slopes of the regression lines is statistically significant ($p < .01$).

SUMMARY

Predictions as to the nature of the regression lines of probability of endorsement on social desirability scale values for High, Middle, and Low *SD* groups were based on an acquiescence and a social desirability interpretation of the *SD* scale. The *SD* scale and a set of 150 items from the EPPS were administered to 237 college students. The subjects were divided into three separate groups on the basis of their scores on the *SD* scale. For each group separately, the regression line of probability of endorsement on social desirability scale values of the EPPS items was obtained. The form of the regression lines was that predicted by the social desirability interpretation of the SD scale, rather than that predicted by the acquiescence interpretation. Furthermore, contrary to the acquiescence interpretation, the mean probability of a True response to the EPPS items was not greater for the Low *SD* group than for the High *SD* group.

TABLE 2

MEANS, VARIANCES, AND REGRESSION COEFFICIENTS OF ITEM PROBABILITIES (Y) AND CORRELATIONS BETWEEN ITEM PROBABILITIES AND SOCIAL DESIRABILITY SCALE VALUES FOR THREE *SD* GROUPS

Group	$\bar{Y}$	s_y^2	r_{xy}	b_y	$s_{y \cdot x}^2$
High *SD*	.645	.086	.844	.28	.025
Middle *SD*	.626	.087	.710	.24	.043
Low *SD*	.633	.069	.731	.22	.032

[2] For the ICL items, the mean probability of endorsement for males was .48 and for females .46.

REFERENCES

COUCH, A., & KENISTON, K. Agreeing response set and social desirability. *J. abnorm. soc. Psychol.*, 1961, **62**, 175–179.

COWEN, E. L., & TONGAS, P. The social desirability of trait descriptive terms: Applications to a self-concept inventory. *J. consult. Psychol.*, 1959, **23**, 361–365.

EDWARDS, A. L. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *J. appl. Psychol.*, 1953, **37**, 90–93.

EDWARDS, A. L. Social desirability and probability of endorsement of items in the Interpersonal Check List. *J. abnorm. soc. Psychol.*, 1957, **55**, 394–396. (a)

EDWARDS, A. L. *The social desirability variable in personality assessment and research*. New York: Dryden, 1957. (b)

EDWARDS, A. L. Social desirability and the description of others. *J. abnorm. soc. Psychol.*, 1959, **59**, 434–436.

EDWARDS, A. L., & WALKER, J. N. Social desirability and agreement response set. *J. abnorm. soc. Psychol.*, 1961, **62**, 180–183.

HANLEY, C. Social desirability and responses to items from three MMPI scales: *D*, *Sc*, and *K*. *J. appl. Psychol.*, 1956, **40**, 324–328.

JACKSON, D. N., & MESSICK, S. Content and style in personality assessment. *Psychol. Bull.*, 1958, **55**, 243–252.

WIGGINS, J. S. Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *J. consult. Psychol.*, 1959, **23**, 419–427.

(Received April 27, 1961)

Psychology and Mental Health

A report of the Institute on Education and Training for Psychological Contributions to Mental Health, held at Stanford University in August, 1955.

Edited by CHARLES R. STROTHER

Topics discussed at the Institute:

Training Needs of Psychologists in Community Mental Health

Specialization in Training

Practicum Training

Training for Therapy

Training for Research in the Mental Health Field

Problems of Departmental Organization

Price: $1.75

Order from:

American Psychological Association, Inc.
1333 Sixteenth Street, N. W.
Washington 6, D. C.

"The new in tests. . . . . ."

Need an individual test which *quickly* provides a stable and reliable estimate of intelligence in 2 or 3 minutes per form? Has three forms? Does not depend on verbal production or subjective scoring? Can be used with the severely physically handicapped (even paralyzed), if they can signal yes-no? Handles two-year-olds and superior adults within the same short series of items and the same format? Has been painstakingly standardized? Builds rapport with intrinsically interesting and simple procedures based on multiple drawings?

Ammons and Ammons

QUICK TEST (QT)

Order Tester's Set including manual, card materials, keys, and 100 answer sheets ($8.00 complete).

•

For further information, write to

PSYCHOLOGICAL TEST SPECIALISTS
Box 1441
Missoula, Montana

"The new in tests. . . . . ."

NEW TEST!!

NEW TEST!!